



**A University of Sussex PhD thesis**

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details



**Investigating the role of Epstein-Barr virus lytic  
key regulator protein Zta in transcriptional  
regulation**

**By**

**Rajaei Almohammed**

**A Thesis submitted for the degree of Doctor of  
Philosophy**

**School of Life Sciences  
University of Sussex**

**September 2016**

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature: .....

## Acknowledgements

First and foremost, I would like to take the opportunity to thank my supervisor Professor Alison Sinclair for her invaluable and continuous support during this project. Alison has always been there for us and did not spare any effort to help us even when she was going through the hardest times in life. I am truly indebted for her patience, guidance and making me feel welcomed as a valuable member of her group.

I would also like to thank my co-supervisor Professor Michelle West for her help and for making a great effort to provide constructive feedback and insightful comments on my work.

Many thanks go to our lab members for their tremendous support. In particular, Dr. Kay Osborn and Dr. Sharada Ramasubramanyan for sharing their knowledge and expertise with me. I would also like to thank Chris Traylen, Barak Perez-Fernandez, Anja Godfrey, and Yaqi Zhou for being my best friends and colleagues at the same time. Besides our lab members, I would like to thank every member of the West's group for their continuous encouragement.

I would like to thank all the undergraduate and master students who undertook their projects in our lab during my Ph.D. project. Being able to share my knowledge with them and observe their success had always been the source of my joy and inspiration.

I am so grateful to my scholarship donor, King's Abdullah scholarship program, who offered me with generosity the opportunity to do a Ph.D. degree.

Last but not least, I would like to thank all my family for their endless support and understanding. Words cannot express how much I miss the feeling of being home in their presence. Many special thanks also go to Deema Jallad; I am truly indebted to her for all the help and thoughtful words of encouragement.

UNIVERSITY OF SUSSEX

RAJAEI ALMOHAMMED

Ph.D. in BIOCHEMISTRY

Investigating the role of Epstein-Barr virus lytic key regulator protein Zta in  
transcriptional regulation

## Summary

Epstein-Barr virus (EBV) is a human herpes virus that, upon primary infection, establishes life-long persistence in B cells (latency). One viral protein, the immediate early lytic protein Zta (also known as BZLF1, ZEBRA, EB1, and Z) plays a significant role in disturbing this latency and inducing a viral productive (lytic) cycle. Expression of Zta, a basic leucine zipper transcription factor, induces a cascade of viral lytic cycle gene expression. The activation of many lytic genes requires the direct binding of Zta to its response elements (ZREs) within proximal promoters of these genes. Much research in recent years has focused on investigating Zta reactivation of latency and its role in lytic viral DNA replication. This has revealed a wealth of knowledge about Zta as a multifunctional transcription factor. However, a complete understanding of Zta transcriptional activation is still missing. Here, utilising ChIP-qPCR, we showed conserved binding patterns for Zta across several EBV lytic gene promoters in two different EBV systems including a non-B cell EBV-infected cell line. Also, using luciferase reporter assays, we show the first functional evidence for a possible role of Zta in controlling transcription regulation at distal regulatory elements (enhancers). Importantly, we identified *BNLF2a*, an essential viral immune evasion gene, as a direct target for Zta. We identified five ZREs and mapped functional ones within the *BNLF2a* promoter using mutational analysis and luciferase reporter assays. We also expressed and purified a recombinant GFP-bZIP Zta protein to address Zta binding to *BNLF2a* ZREs *in vitro*. Interestingly, using *in silico* approach, we also identified a conserved sequence in the ZRE flanking region of all five ZREs within *BNLF2a* promoter and uncovered a role for a possible repressor at ZRE2 flanking region. Our work not only adds to our understanding of Zta transcription regulation but characterises for the first time the regulation of a novel Zta target that has a role in evading the host immune system during EBV pre-latency and lytic cycle.

# Table of Contents

<b>Chapter 1 Introduction</b>	<b>1-14</b>
<b>1.1 Viruses and cancer</b>	<b>1-14</b>
<b>1.2 Herpesviruses</b>	<b>1-18</b>
<b>1.3 Epstein- Barr virus (EBV)</b>	<b>1-20</b>
1.3.1 EBV structure	1-20
1.3.2 EBV genome	1-23
1.3.3 EBV strains and subtypes	1-27
1.3.4 EBV life cycle	1-28
1.3.5 EBV cell tropism	1-44
1.3.6 EBV-associated diseases	1-46
<b>1.4 Zta (BZLF1)</b>	<b>1-60</b>
1.4.1 Zta structure and DNA binding	1-61
1.4.2 Zta promoter (Zp)	1-65
1.4.3 Zta protein interactions and functional diversity	1-67
1.4.4 Zta transcriptional regulation function	1-69
<b>1.5 EBV immune evasion</b>	<b>1-71</b>
1.5.1 The role of BNLF2a in immune evasion	1-73
<b>1.6 Project Aims</b>	<b>1-75</b>
<b>Chapter 2 Materials and Methods</b>	<b>2-76</b>
<b>2.1 Materials, and reagents</b>	<b>2-76</b>
2.1.1 DNA constructs (plasmids)	2-76
2.1.2 Cell lines	2-77
2.1.3 Antibodies	2-77
2.1.4 Primers (oligonucleotides)	2-78

2.1.5 Solutions and buffers	2-80
2.1.6 Kits and reagents	2-82
<b>2.2 Methods</b>	<b>2-84</b>
2.2.1 General nucleic acid methods	2-84
2.2.2 General protein methods	2-92
2.2.3 Tissue culture	2-93
2.2.4 Luciferase reporter assays	2-95
2.2.5 Chromatin immunoprecipitation (ChIP-qPCR)	2-96
2.2.6 Protein expression and purification (His-GFP-bZIP Zta)	2-99
2.2.7 HeLa cell nuclear extract	2-101
2.2.8 Electrophoretic mobility shift assay	2-101
2.2.9 Computational methods	2-103
<b>Chapter 3 Zta binding across the viral genome</b>	<b>3-104</b>
3.1 Introduction	3-104
3.2 Results	3-106
3.2.1 ChIP-qPCR approach to analyse Zta binding to EBV genome	3-106
3.2.2 Zta binds similarly to EBV OriLyt	3-112
3.2.3 Zta binds similarly to promoter regions of the immediate-early genes	3-114
3.2.4 Zta binds similarly to promoter regions of various early lytic genes	3-116
3.2.5 Zta binds similarly to promoter regions of various late lytic genes	3-118
3.2.6 Zta binds to a novel site with unknown regulatory function	3-120
3.3 Discussion	3-124
<b>Chapter 4 ZREs as long-distance enhancer elements</b>	<b>4-129</b>
4.1 Introduction	4-129
4.2 Results	4-132
4.2.1 Luciferase reporter assay to investigate ZREs as potential enhancer elements	4-132

4.2.2 Zta activates a heterologous minimal promoter through distal ZREs-----	4-136
4.2.3 Zta activates a known target promoter through distal ZREs -----	4-138
<b>4.3 Discussion -----</b>	<b>4-140</b>
<b><i>Chapter 5 Zta activates viral immune evasion gene BNLF2a through direct binding to proximal promoter ZREs-----</i></b>	<b><i>5-143</i></b>
<b>5.1 Introduction -----</b>	<b>5-143</b>
<b>5.2 Results -----</b>	<b>5-146</b>
5.2.1 Zta binds at the BNLF2a promoter during lytic cycle -----	5-146
5.2.2 BNLF2a predicted ZRE sequences.-----	5-150
5.2.3 BNLF2a promoter constructs to investigate each ZRE function -----	5-152
5.2.4 Zta activates BNLF2a through promoter ZREs in different cell lines -----	5-159
5.2.5 The proximal ZREs to the TSS are essential for BNLF2a activation.-----	5-161
5.2.6 ZREs functional redundancy-----	5-165
5.2.7 In vitro analysis of Zta binding to BNLF2a ZREs -----	5-172
<b>5.3 Discussion -----</b>	<b>5-180</b>
<b><i>Chapter 6 Other elements play a role in BNLF2a gene regulation -----</i></b>	<b><i>6-187</i></b>
<b>6.1 Introduction -----</b>	<b>6-187</b>
<b>6.2 Results -----</b>	<b>6-189</b>
6.2.1 The conservation of a motif flanking BNLF2a ZREs -----	6-189
6.2.2 Investigating other regulatory elements within BNLF2a promoter-----	6-191
6.2.3 A single point mutation in the TATA box enhances Zta activation of BNLF2a -----	6-197
6.2.4 The KLF4 site plays a role in BNLF2a promoter regulation-----	6-201
6.2.5 Mutational analysis of the motif flanking ZREs revealed a repressor binding site -----	6-204
6.2.6 Zta binding to ZRE2 is not affected by mutation in the flanking motif -----	6-209
6.2.7 Specific complex binds to ZRE2 flanking sequence (E-box) in HeLa cells nuclear extract -	6-211



6.3 Discussion .....	6-213
<b><i>Chapter 7 General discussion</i></b> .....	<b>7-219</b>
<b><i>References</i></b> .....	<b>7-225</b>
<b><i>Appendix A. Publications</i></b> .....	<b>7-255</b>

## Table of Figures

Figure 1-1 EBV structure.....	1-22
Figure 1-2 EBV genome .....	1-26
Figure 1-3 EBV life cycle .....	1-32
Figure 1-4 Zta (BZLF1) structure .....	1-64
Figure 1-5 BZLF1 locus and promoter .....	1-66
Figure 3-1 Differences in Zta binding to the EBV genome. ....	3-109
Figure 3-2 Experimental design of Zta ChIP-qPCR.....	3-111
Figure 3-3 Zta binding at the origin of lytic replication (OriLyt). ....	3-113
Figure 3-4 Zta binding at the immediate-early lytic promoters.....	3-115
Figure 3-5 Zta binding at the early lytic promoters.....	3-117
Figure 3-6 Zta binding at the late lytic promoters.....	3-119
Figure 3-7 Zta binding at a region with unknown regulatory function (~82kbp). ....	3-122
Figure 3-8 Sequence variations between EBV strain at the ~82kbp region. ....	3-123
Figure 4-1 Investigating a long-range ZRE element effect on promoter activation. ....	4-134
Figure 4-2 DNA sequences of the long-range ZREs element and heterologous promoters.....	4-135
Figure 4-3 A long-range ZRE element drives Zta activation of a heterologous minimal promoter. ..	4-137
Figure 4-4 A long-range ZRE element drives Zta activation of a known Zta target promoter. ....	4-139
Figure 4-5 Increasing evidence suggesting a role for Zta at enhancer elements. ....	4-142
Figure 5-1 Zta binding at BNLF2a promoter.....	5-148
Figure 5-2 Zta binding at BNLF2a promoter region in different cell lines. ....	5-149
Figure 5-3 Zta response elements (ZREs) upstream of BNLF2a gene. ....	5-151
Figure 5-4 Investigating BNLF2a promoter activity. ....	5-154
Figure 5-5 BNLF2a promoter sequence along with the predicted ZRE motifs. ....	5-155
Figure 5-6 A schematic diagram of BNLF2a promoter.....	5-156
Figure 5-7 BNLF2a ZREs luciferase promoter constructs. ....	5-158
Figure 5-8 Zta activates BNLF2a through ZREs.....	5-160
Figure 5-9 The proximal ZREs contribute the most to BNLF2a activity. ....	5-164

Figure 5-10 Mutating one ZRE of the proximal ZREs has no profound effect on promoter activity. .	5-169
Figure 5-11 The effect of leaving one ZRE of the first three proximal ZREs on promoter activity. ....	5-170
Figure 5-12 The effect of leaving two ZRE of the proximal ZREs on promoter activity. ....	5-171
Figure 5-13 Generating His-GFP-bZIP Zta expression vector. ....	5-174
Figure 5-14 His-GFP-bZIP Zta sequences. ....	5-175
Figure 5-15 Large scale purification of His-GFP-bZIP Zta protein. ....	5-176
Figure 5-16 Purified His-GFP-bZIP Zta. ....	5-177
Figure 5-17 Interaction of Zta with BNLF2a ZREs in vitro. ....	5-179
Figure 5-18 The ZREs contribution to the promoter activity. ....	5-185
Figure 5-19 A fail-safe mechanism for BNLF2a promoter activation. ....	5-186
Figure 6-1 Position weight matrices (PWMs) of a conserved ZRE flanking motif. ....	6-190
Figure 6-2 BNLF2a promoter and the predicted ZREs flanking motif sequence.....	6-192
Figure 6-3 BNLF2a promoter wildtype and mutant elements. ....	6-193
Figure 6-4 A schematic diagram of BNLF2a promoter showing the ZREs and other elements. ....	6-194
Figure 6-5 BNLF2a luciferase promoter constructs. ....	6-196
Figure 6-6 BNLF2a activation in the presence of a non-canonical TATA-box in DG75 cells.....	6-199
Figure 6-7 BNLF2a activation in the presence of a non-canonical TATA-box in HeLa cells.....	6-200
Figure 6-8 KLF4 binding site has an effect on basal promoter activity. ....	6-203
Figure 6-9 ZREs flanking motifs effects on Zta activation of BNLF2a promoter. ....	6-206
Figure 6-10 ZRE2 flanking motif has a profound effect on basal promoter activity. ....	6-208
Figure 6-11 Interaction of Zta with the mutant motif flanking ZRE2. ....	6-210
Figure 6-12 EMSA using HeLa nuclear extract.....	6-212
Figure 6-13 A model for the possible role of ZRE2 flanking motif (E-box). ....	6-218

## Table of Tables

Table 1.1 Cancers caused by viruses. ....	1-17
Table 1.2: Human herpesviruses. ....	1-20
Table 1.3 Latency program. ....	1-36
Table 2.1: Plasmid DNA constructs used in various experiments. ....	2-77
Table 2.2: Mammalian cell lines used in various experiments. ....	2-77
Table 2.3: Antibodies used in various experiments. ....	2-78
Table 2.4: Primers used in ChIP-qPCR experiments. ChIP-qPCR for EBV target genes. ....	2-79
Table 2.5: Oligonucleotides used to generate EMSA probes. ....	2-80
Table 2.6: Solutions and buffers. ....	2-81
Table 2.7: Kits and reagents. ....	2-84

## Abbreviations

Abbreviation	Meaning
3C	Chromosome conformation capture
AP1	Activator protein 1
ATF	Activating transcription factor
BARTs	BamHI A rightward transcript
BCR	B-cell receptor
BL	Burkitt's Lymphoma
bp	Base pair
bZIP	Basic leucine zipper
C/EBP	CCAAT/enhancer binding protein
CBP	CREB binding protein
ChIP	Chromatin immunoprecipitation
cHL	Classical Hodgkin lymphoma
CREB	cAMP response element-binding protein
CTLs	Cytotoxic lymphocytes
DBD	DNA-binding domain
DIM	Dimerisation domain
EBER	Epstein–Barr virus-encoded small RNAs
EBNA	Epstein-Barr nuclear antigen
EBV	Epstein-Barr virus
EMSA	Electrophoretic mobility shift assay
EP300	E1A-binding protein p300
ESCs	Embryonic stem cells
FBS	Foetal bovine serum
HL	Hodgkin Lymphoma
IL	Interleukin
kbp	Kilobase pair
KDa	Kilodalton
KLF4	Kruppel like factor 4
KSHV	Kaposi's sarcoma-associated herpesvirus
LB	Luria broth

LCL	Lymphoblastoid cell line
LMP	Latent membrane protein
MAP	mitogen-activated protein
MHC	Major histocompatibility complex
miRNA	Micro RNA
mRNA	Messenger RNA
NE	Nuclear
NF $\kappa$ B	Nuclear factor kappa enhancer of activated B cell
NK	Natural killer
NPC	Nasopharyngeal carcinoma
nt	Nucleotides
OHL	Oral hairy leukoplakia
OriP	Origin of plasmid replication
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PKC	Protein kinase C
Pol II	RNA polymerase II
PTLD	Post-transplant lymphoproliferative diseases
qPCR	Quantitative PCR
RE	Restriction enzyme
Rp	Rta (BRLF1) promoter
RRE	Rta response element
SCC	Squamous cell carcinoma
SP1	Specificity protein 1
TBP	TATA box binding protein
TPA	Tetradecanoyl phorbol acetate
TSS	Transcription start site
UTR	Untranslated region
WB	Western blot
ZEB	Zinc finger E-box-binding homeobox
Zp	Zta (BZLF1) promoter
ZRE	Zta response element

## **Chapter 1 Introduction**

### **1.1 Viruses and cancer**

Viruses were initially described as filterable agents and were known to be associated with diseases as early as 1898. In this year, a Dutch scientist, Martinus Beijerinck not only replicated previous work showing that a filterable agent (which he called a “contagium vivum fluidum”) was capable of inducing disease in tobacco plant but more importantly demonstrated the fact that this agent gained potency after replicating in plant cells (Fields et al., 2013). Beijerinck’s findings clearly indicated that this filterable pathogen which was smaller than bacteria reproduced itself in living cells in contrast to an earlier hypothesis by the Russian scientist Dimitri Ivanofsky, who suggested that this filterable material could be a toxin (Fields et al., 2013).

Around the same time, two German scientists, Friedrich Loeffler and Paul Frosch, presented the first evidence that pathogens smaller than bacteria are capable of causing diseases in animals. In their experiments, they were able to show the transmission of the foot and mouth infection from infected animals to healthy livestock by using only germ-free filtered vesicle lymph. What’s more, they illustrated that animals injected with infectious vesicle lymph that had been heated, developed immunity to the disease. This led to their conclusion that an ultra-filterable and ultra-visible particle is the cause of the foot and mouth disease in animals (Rott and Siddell, 1998).

These pioneering findings were soon followed by other major discoveries of viruses that cause devastating diseases in humans such as the yellow fever virus described by Walter Reed in 1901 (Reed et al., 1901) and poliovirus described by Karl Landsteiner and Erwin Popper in 1909 (De Jesus, 2007)).

However, the link between cancer and viruses was not established until the early 20th century. In 1908, two Danish scientists, Ellerman and Bang demonstrated that a filterable agent could transmit leukaemia in chickens (Cardiff and Kenney, 2007). Three years later, Peyton Rous was also able to

show the transmission of chicken sarcoma tumour by injecting healthy chickens with cell-free tumour filtrate. Rous reported his conclusion that this tumour is transferable through a small filterable agent, “The first tendency will be to regard the self-perpetuating agent active in this sarcoma of the fowl as a minute parasitic organism” (Rous, 1911).

Despite this early evidence that viruses cause cancer in animals, the origin of these tumors remained highly debatable, and many pioneer scientists were sceptical about the emerging theories that viruses could cause cancer in humans.

It took nearly another 50 years before the first evidence of a human tumour virus was reported. A fruitful collaboration between two British doctors Denis Burkitt and Anthony Epstein led to the identification of the first human tumour virus, a new herpesvirus (known today as Epstein-Barr virus (EBV) or human herpesvirus 4).

In 1958, Denis Burkitt, who was working in Uganda at that time, encountered many cases of a very unusual lymphoma that was seen in young children with a high incidence rate (Burkitt, 1958). Burkitt mapped the geographical region where high rates of this lymphoma were often reported. The geographic distribution of the tumour was mostly in sub-Saharan areas where temperature and rainfall are usually high (Burkitt, 1962a, b; Burkitt and Wright, 1963; Burkitt, 1961). Interestingly, this geographic distribution correlated with the distribution of the malaria parasite, *Plasmodium falciparum* (Dalldorf et al., 1964; Guerra et al., 2008). Indeed, malaria was later shown to play a role in the development of Burkitt’s lymphoma as will be discussed in **Section 1.3.6.2**.

While visiting London to present his findings at the Middlesex Hospital Medical School, Burkitt met Anthony Epstein, who was at the time working as a researcher at this hospital. On viewing data presented by Burkitt, Epstein was immediately convinced that a pathogen might be causing this lymphoma. He further speculated that this pathogen could be a tumour virus transmitted by an arthropod (Crawford, 2014).



Epstein had gained significant experience in electron microscopy and tumour viruses from his previous work on Rous fowl sarcoma virus (Epstein and Holt, 1958) and he knew that to prove his hypothesis, he needed to show the presence of a virus in these lymphoma cells. Luckily, it took Epstein only a few years to prove his theory. In 1964, Anthony Epstein and his colleagues Yvonne Barr and Bert Achong were able to show herpes virus-like particles in cell lines derived from Burkitt's lymphoma using electron microscopy (Epstein et al., 1964). However, this did not prove that the newly identified virus was the cause of this lymphoma. Epstein therefore published his work without concluding any link between the virus and Burkitt's lymphoma (Epstein et al., 1964).

Fortunately for the scientific community, this was not the end of the story. Indeed, further links between EBV and cancer were revealed a few years later through another successful collaboration between Epstein and the Werner Henle group in Philadelphia, USA (Epstein et al., 1965). The Henle group were able to show immunofluorescent staining in the lymphoma cell lines, using various sera from Burkitt's lymphoma patients that were thought to be specific to an EBV antigen in these cells (Henle and Henle, 1966).

The Henle group were also able to establish the first concrete evidence of EBV lymphoproliferative effects by co-culturing healthy B cells with lethally irradiated EBV-positive Burkitt's lymphoma cells, which induced the healthy B cells to grow in culture for a few weeks (Henle et al., 1967). Soon after, they were able to establish continuously proliferating lymphoblastoid cell lines (LCLs) from EBV umbilical cord lymphocytes co-cultured with EBV Burkitt's cell lines (Nilsson et al., 1971). The scientific and personal events that led to the discovery of EBV have been the subject of several reviews (Crawford, 2014; Epstein, 2015; Epstein, 2001; Young and Murray, 2003).

Currently, cancers associated with virus infections are estimated to account for nearly 15% of all cancer cases worldwide (Moore and Chang, 2010). There are seven viruses classified as definitely carcinogenic to human. These are Epstein-Barr virus (human herpesvirus 4), Kaposi's sarcoma herpesvirus (human

herpesvirus 8), hepatitis B and C viruses, human papillomavirus (HPV), human T-cell lymphotropic virus, and Merkel cell polyomavirus (**Table 1.1**) (Oh and Weiderpass, 2014).

Despite the fact that this percentage of cancer cases is considered relatively low the high prevalence of these viruses in some cancers cannot be ignored. For instance, in endemic Burkitt's lymphoma and nasopharyngeal carcinoma, EBV is found in nearly 100% of the cases in high incidence regions. Similarly, high HPV prevalence is seen in most cervix carcinomas (de Martel et al., 2012).

<b>Virus</b>	<b>Associated cancers</b>
Epstein-Barr virus	Nasopharyngeal carcinoma, Burkitt's lymphoma, non-/Hodgkin lymphoma and gastric carcinoma
Hepatitis B and C virus	Hepatocellular carcinoma
Kaposi sarcoma herpes virus	Kaposi sarcoma
Human immunodeficiency virus type 1	Kaposi sarcoma, non-/Hodgkin lymphoma, carcinoma of the cervix, anus and conjunctiva
Human papillomavirus type 16	Carcinoma of the cervix, vulva, vagina, penis, anus, oral cavity, and oropharynx and tonsil
Human T-cell lymphotropic virus type 1	Adult T-cell leukaemia and lymphoma

**Table 1.1 Cancers caused by viruses.**

It is worth noting that viruses are not only seen as a cause of cancer but, possibly, a way to prevent and cure cancer. For example, vaccine development for tumour viruses will help to reduce cancer cases. This is clearly seen in the development of human papillomavirus (HPV) vaccine, which has been used in preventing cervical cancer in women (Crosignani et al., 2013).

Furthermore, oncolytic viruses are currently leading the way in viral cancer therapy. Such viruses are being engineered to infect and replicate within specific tumour cells; this replication leads to tumour cell death. Furthermore, these viruses provoke a host immune system response against the tumour cells which will further help in destroying these cells (Kaufman et al., 2015). Indeed, a recently FDA approved modified herpes simplex virus-1 (HSV-1) has shown

promising results in treating advanced melanoma (Andtbacka et al., 2015; Emens et al., 2016).

These facts highlight the importance of studying tumour-virus biology to understand the development of many viral related cancers so that new prevention and treatment measures can be developed.

## 1.2 Herpesviruses

Herpesviruses are double-stranded DNA viruses which are one of the most significant pathogens associated with tumors and several diseases in human as well as other hosts. Clear examples of such diseases are genital herpes caused by herpes simplex viruses, infectious mononucleosis caused by EBV, and Kaposi's sarcoma caused by Kaposi's sarcoma-associated herpes virus (KSHV) (Arvin et al., 2007). Many of these diseases were known to humans for centuries. Indeed, the word herpes, which means "to creep" in Greek, was used as a medical term to describe lesions in ancient Greece times 2500 years ago (Beswick, 1962).

Viruses in the herpesviruses family (*Herpesviridea*) were initially classified based on their shared biological features such as the primary natural host, viral structure and genome (Roizman et al., 1981). However, more recently, the taxonomy of *herpesviridae* was revised by the International Committee on Taxonomy of Viruses (ICTV) to the new order *Herpesvirales* (Davison et al., 2009). This new order is split into three families: *Herpesviridae* which retains the mammal, bird and reptile viruses, *Alloherpesviridae* which comprises the fish and frog viruses, and *Malacoherpesviridae* which consists of a bivalve virus. A new genus incorporating new species has also been added under the subfamilies of the *Herpesviridae* family (Davison, 2010; Davison et al., 2009).

The *Herpesviridae* family, which includes the mammalian herpesviruses is divided into three subfamilies: alpha-, beta- and gamma- herpesvirinae. This classification was initially based on structural similarities (Roizman et al., 1981). However, more advanced phylogenetic tools have not resulted in any changes to this classification (McGeoch et al., 1995). These subfamilies comprise nine

known human herpesviruses (listed in **Table 1-2**). The alphaherpesviruses include the herpes simplex viruses (HSVs) and the varicella-zoster virus (VZV). These viruses establish latency primarily in sensory ganglia but can infect a variety of cells. The betaherpesviruses include cytomegalovirus which can infect a variety of human lymphocytes but mostly cells of myeloid origins (Arvin et al., 2007).

The *gammaherpesviruses*, which infect lymphocytes and induce lymphoproliferation, is further split into two genera: the *lymphocryptoviridae* (LCV) and *rhadinoviridae* (RV). The *lymphocryptoviridae* herpesviruses contain Epstein-Barr virus (EBV) and other related viruses identified in various primates species while, the *rhadinoviridae* herpesviruses include Kaposi's sarcoma-associated herpes virus (KSHV) (known as human herpesviruses 8 (HHV-8)) (Arvin et al., 2007).

Name	Other given names	Genome size (kb)	<i>Herpesviridae</i> subfamily
Human herpesvirus 1 (HHV-1)	Herpes simplex virus 1 (HSV-1)	152	<i>Alphaherpesvirinae</i>
Human herpesvirus 2 (HHV-2)	Herpes simplex virus 2 (HSV-2)	155	<i>Alphaherpesvirinae</i>
Human herpesvirus 3 (HHV-3)	Varicella-zoster virus (VZV)	125	<i>Alphaherpesvirinae</i>
Human herpesvirus 4 (HHV-4)	Epstein-Barr virus (EBV)	172	<i>Gammaherpesvirinae</i> : <i>Lymphocryptovirus</i> genera
Human herpesvirus 5 (HHV-5)	Cytomegalovirus (CMV)	236	<i>Betaherpesvirinae</i>
Human herpesvirus 6A (HHV-6A)	HHV-6 variant A	159	<i>Betaherpesvirinae</i>
Human herpesvirus	HHV-6 variant B	162	<i>Betaherpesvirinae</i>

6B (HHV-6B)			
Human herpesvirus 7 (HHV-7)		145	<i>Betaherpesvirinae</i>
Human herpesvirus 8 (HHV-8)	Kaposi's sarcoma- associated herpesvirus (KSHV)	170	<i>Gammaherpesvirinae:</i> <i>Rhadinoviridae</i> genera

**Table 1.2: Human herpesviruses.**

In herpesviruses, it is believed that, apart from the fish herpesviruses (*Alloherpesviridae*), all mammalian and avian herpesviruses share a common origin (McGeoch et al., 1995). For the mammalian herpesviruses (*Herpesviridae*), the three subfamilies (alpha, beta, and gamma) appear to share a common root and may have appeared around 180 to 220 million years ago, while the major sublineages under these subfamilies may have appeared 80 to 60 million years ago (McGeoch et al., 1995).

### 1.3 Epstein- Barr virus (EBV)

The gammaherpesvirus EBV is a DNA virus that infects over 90% of the population worldwide (Borer et al., 1999; de-The et al., 1975) EBV establishes a life-long persistence in B cells and can also infect epithelial cells.

#### 1.3.1 EBV structure

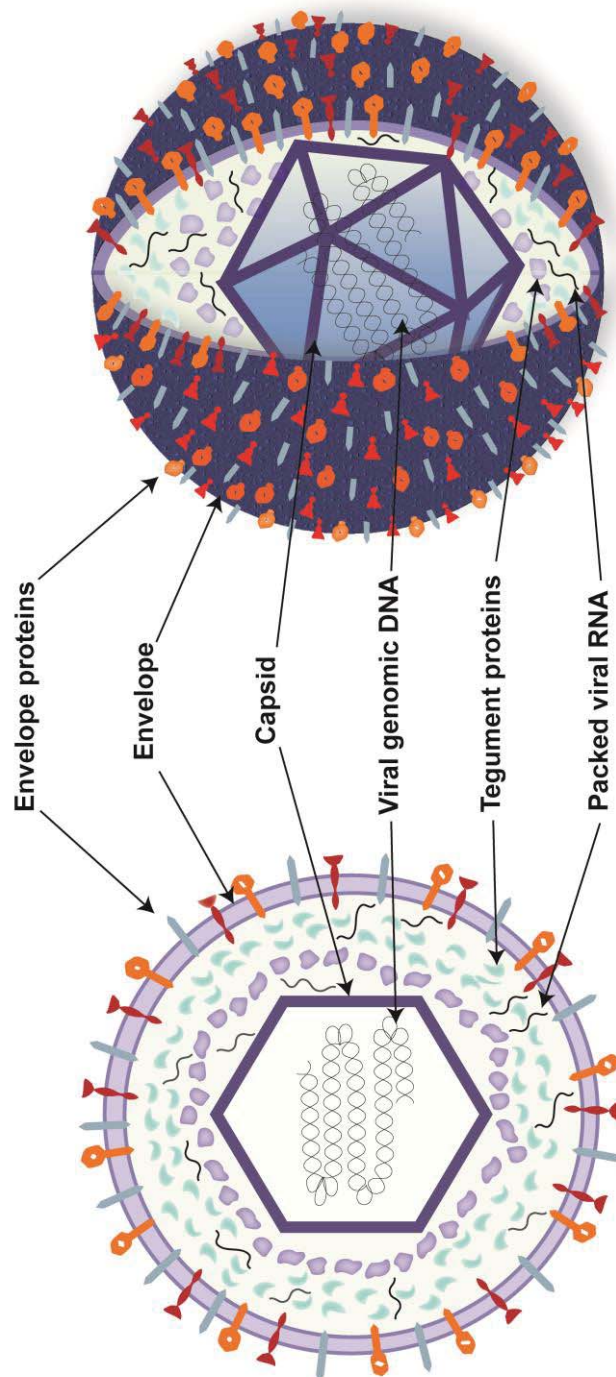
In common with other herpesviruses, the EBV virion ranges in size between 120-260 nm and has a DNA core containing a single molecule of linearised naked viral double-stranded DNA (Pellett and Roizman, 2013).

As illustrated in **Figure 1-1**, this DNA is contained in a nucleocapsid which consists of 162 subunits called capsomeres. The nucleocapsid is composed of several proteins including major capsid and minor capsid proteins. The capsid is separated from the outer envelope by tegument proteins. These various proteins are distributed asymmetrically and are believed to play a significant

role in mediating the early events of viral entry. Multiple late lytic genes encode for tegument proteins such as *BPLF1*, *BOLF1*, *BGLF4*, *BRRF2*, *BKRF4* and *BNRF1* (Longnecker et al., 2013; Pellett and Roizman, 2013).

More recently, a study by (Jochum et al., 2012) showed the existence of various packed viral RNAs within the virion. These RNAs include viral lytic transcripts such as viral *BZLF1* and *BNLF2a* transcripts, miRNAs, and other noncoding RNAs. The virion-packed viral RNAs role appears to be important in triggering a pre-latent EBV infection phase, inducing cellular cytokines, and host immune evasion, as is known for the *BNLF2a* transcript (Jochum et al., 2012).

The outer envelope is composed of external glycoprotein spikes; there are at least nine different glycoproteins in EBV envelope including gp350, gp150, gH, gB and gL (Longnecker et al., 2013; Pellett and Roizman, 2013).



**Figure 1-1 EBV structure.** EBV virion consists of a double stranded DNA surrounded by icosahedral capsid. The capsid is enclosed by a lipid envelope with viral glycoproteins spikes that are used to attach and infect the host cell. The area between the capsid and the envelope contains teguments proteins and packed viral mRNA. Image adapted from (Longnecker et al., 2013).



### 1.3.2 EBV genome

The EBV genome is double-stranded DNA, about 172 kbp long, and has 60% of GC content. The viral DNA exists as either linear packed within the virion or circular in infected cells. The linearised genome becomes circular episome inside the host cell by the fusion of the two-terminal repeat ends of the linear DNA. The EBV genome contains nearly 84 protein-coding open reading frames (ORF) (Longnecker et al., 2013). In addition to this, EBV encodes for various noncoding RNAs. The nonpolyadenylated EBV-encoded small RNAs EBER (1&2) (Lerner et al., 1981), BamHI-A rightward transcripts (BARTs), and miRNAs transcribed from within BHRF1 ORF and BARTs locus (Pfeffer et al., 2005; Pfeffer et al., 2004).

As illustrated in **Figure 1-2**, the genome consists of a unique long sequence containing coding regions interspersed by four major internal repeats (IR1 to IR4). The separate coding regions resemble five unique segments referred to as U1-U5. The length of each internal repeat is different: IR1 is the longest (3072 bp), IR2 is 123 bp, IR3 is 708 bp, and IR4 is 103 bp. The two tandem repeat regions at each end of the linear genome (terminal repeats (TR)) are around 500 bp (Longnecker et al., 2013).

IR1, the longest internal repeat, consists of similarly arranged repeats (3 kbp in length) known as BamHI W repeats. The repeat contains a promoter (Wp), and two exons referred to as W1 and W2 coding for the repeats in latent EBNA-LP protein, as will be discussed in subsequent sections. The number of the W repeats is variable between different EBV isolates (Allan and Rowe, 1989). IR2 and IR4 (also known as DL and DR, respectively) share highly similar sequence (Dambaugh and Kieff, 1982). IR3 is characterised by tandem repeats of a GGA motif and is part of *BKRF1* gene, which encodes for the latent protein EBNA1 (Heller et al., 1985; Heller et al., 1982).

The first complete sequencing of the EBV isolate B95-8 was achieved by sequencing clones from EBV DNA BamHI-digested fragments (Baer et al., 1984). In this B95-8 sequence, the first nucleotide was considered to be the base located before the leftmost EcoRI restriction site within the U1 coding



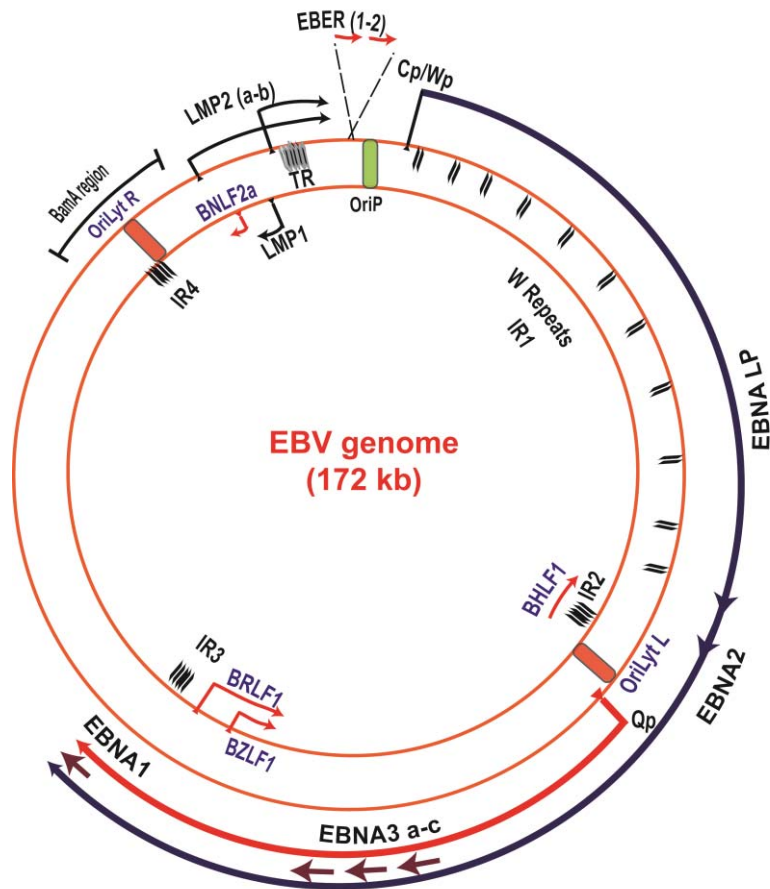
region. The nomenclature of the various ORFs was based on the relative location and transcription orientation of the ORF within the DNA fragments that resulted from BamHI restriction digest of the viral DNA. The BamHI fragments were given alphabetical letters (A-Z) where A is the largest fragment and Z is the smallest. For example, an ORF such as BZLF1 will be the first ORF within BamHI Z (BZ) fragment and has a leftward orientation (LF). If the ORF is located over two BamHI fragments, the name is based on the 5' fragment (Longnecker et al., 2013; Sample et al., 2009). In the published literature, most lytic cycle genes are referred to by their gene names. However, latent cycle genes are commonly known by their protein names.

The EBV genome has a single origin of replication during latency (OriP). This origin of EBV DNA replication is about 1.7 kbp in length and has two functional domains, the family of repeats (FR) and the Dyad symmetry (DS) element. Replication during latency is mediated through EBNA1, which binds to multiple sites on FR and DS, and subsequently recruits the cellular DNA replication machinery (Yates et al., 2000).

Replication during the EBV lytic cycle is not dependent on latency DNA replication. Unlike in latency, the EBV genomic DNA is amplified up to several hundred-fold within 24 hours (Hammerschmidt and Sugden, 1988). During EBV lytic cycle, DNA replication is mediated through two identical copies of two regions across the viral genome separated by nearly 100 kbp. These two regions are known as the origins of replication during EBV lytic cycle; OriLyt Left (L) and OriLyt Right (R). These regions are also referred to as left and right duplicated sequences of the genome (DSL and DSR) (Hammerschmidt and Sugden, 1988; Schepers et al., 1993b). Each OriLyt consists of a 1055 bp sequence known as the OriLyt core element. The core element is identical in both regions and consists of two essential replication components separated by approximately 530 bp. The essential components contain multiple binding sites for viral *BZLF1* as well as cellular proteins such as the transcription factors SP1 and ZBP-89. In OriLyt L, *BHLF1* and its promoter is found at one end of the OriLyt core element, while *BHRF1* along with its promoter is found at the other end. These flanking genes are known as auxiliary components and contribute to

enhancing OriLyt activity. The whole region of the OriLyt including the core promoters of *BHLF1* and *BHRF1* genes span around 7.7 kbp (Hammerschmidt and Sugden, 2013).

It is worth noting that a deletion of one OriLyt does not affect the ability of EBV to replicate during lytic cycle. Naturally occurring EBV strains with such deletions exist. For example, the B95-8 (deletion of OriLyt R) and P3HR1 (deletion of OriLyt L) strains have the ability to induce lytic cycle and produce new virions (Cho et al., 1984; Raab-Traub et al., 1980). A diagram illustrating the main features of EBV genome is presented in **Figure 1-2**.

**A****B**

**Figure 1-2 EBV genome.** The EBV genome is about 172 kbp double-stranded DNA. The viral DNA exists as either linear within viral virions **(A)** or circular in infected cells **(B)**. There are four internal repeat regions (IR 1-4) separating the unique coding regions (U1-5) and two terminal repeat regions (TR), which are fused together to give the circular genome **(A)**. OriP (green) is used to replicate the EBV genome during latency. OriLyts R and L (orange) drive DNA replication during lytic cycle. OriLyt R is deleted in EBV B95-8 strain. The immediate-early lytic proteins and the early lytic BNLF2a ORFs are shown in **(A)** as red arrows. Image adapted from (Longnecker et al., 2013; Price and Luftig, 2014).

### 1.3.3 EBV strains and subtypes

Worldwide EBV isolates are grouped into two types, known as type 1 (A) (e.g. B95-8 strain) and type 2 (B) (e.g. AG876 strain). This classification is based on differences in EBNA2 gene sequence. The identity between EBNA2 in the two strains is only 54% at the protein level. Type 1 EBV is the most prevalent type worldwide. However, the distribution of type 2 EBV is equal to type 1 EBV in certain geographical regions such as sub-Saharan Africa or even dominant over type 1 in other areas such as in Mexico (Farrell, 2015). An important feature of Type 1 EBV is that it is more efficient at immortalising B cells *in vitro* than type 2, emphasising the role of EBNA2 in transforming B cells (Rickinson et al., 1987).

As described earlier, the prototypical type 1 EBV strain B95-8 strain was the first available complete sequence of the EBV genome (Baer et al., 1984). This strain, which was isolated from an infectious mononucleosis patient (Blacklow et al., 1971), has a deletion of about 13.6 kbp that spans IR4 and OriLyt R region affecting three open reading frames including LF1 and LF2, (Farrell, 2001; Farrell, 2015; Parker et al., 1990). The sequence of the deleted region has been determined based on EBV Raji strain (Parker et al., 1990).

The available wildtype EBV type 1 sequence is based on B95-8 sequence integrated with the Raji sequence to complete the deleted region. The wildtype strain is the most annotated EBV genome sequence and available through NCBI accession number (NC\_007605.1). In contrast to the B95-8 strain, EBV identified in Raji cells is incapable of complete lytic replication due to various deletions in the reading frames for its structural proteins (Hatfull et al., 1988). As mentioned earlier, another deletion (6.6-8.5-kbp) affecting OriLyt L, is also reported in various Burkitt's lymphoma EBV isolates such as in Daudi and P3HR1 (Farrell, 2015).

With the advancement of second generation sequencing technologies, many other strains from different parts of the world are being sequenced. Interestingly, particular EBV strains are more prevalent in certain tumors and

ethnic groups which may suggest a strong link between certain EBV strains and the tumor development. For example, the EBV type 2 strain (AG876) has been sequenced (NCBI accession number DQ279927) (Dolan et al., 2006). Also, complete DNA sequences are now available for many EBV isolates from different BL cells such as Raji, Akata, and Mutu. Furthermore, genome sequences of EBV strains such as GD1, GD2, HKNPC1, and M81, isolated from nasopharyngeal carcinoma Chinese patients, are available (Farrell, 2015; Kwok et al., 2012; Tsai et al., 2013).

It is worth noting that sequence polymorphisms have been documented in various EBV latent genes such as in LMP1, LMP2A, EBNA1, and EBNA3B. EBV latency and latent genes will be discussed in the next sections. Apart from the known variations in latent EBNA2 between EBV type1 and type 2, there is either little evidence or more investigations are needed to establish any functional relevance of these variations (Farrell, 2015).

Similarly, such sequence variations are seen in lytic cycle genes such as in *BZLF1*, *BRLF1*, *BHRF1* and *BNLF2a* (Farrell, 2015). In particular, three *BZLF1* promoter (Zp) natural existing variants were shown to be associated with different types of lymphoma (Lorenzetti et al., 2014; Martini et al., 2007).

### **1.3.4 EBV life cycle**

#### **1.3.4.1 EBV biphasic life cycle**

In common with other herpesviruses, EBV shows a biphasic life: A dormant life phase known as latency and an acute productive phase (lytic cycle). Evidence for the EBV biphasic life cycle was observed initially in the differences between EBV producer and non-producer cell cultures. EBV viral capsid antigen (VCA) and early antigen (EA), which are specific for EBV lytic cycle, are detected in the majority of EBV producer cell lines. On the contrary, these antigens were barely detected in non-producer cell lines. Also, EBNA1, a protein expressed during latency, was detected in over 90% of non-producer cells (Klein et al., 1972; Reedman and Klein, 1973).

In latency, the virus establishes a life-long persistence in B cells as an episomal DNA, and only expresses a limited set of genes. On the contrary, during lytic cycle, the virus initiates a cascade of gene expression that facilitates the replication and packaging of EBV DNA into newly produced virus particles (virions).

The EBV latency phase can be seen in healthy memory B cells, lymphomas and cancers from epithelial origins; while the lytic phase is typically seen in permissive epithelial cells and primary infected B cells (pre-latency). Latency is essential for EBV genome persistence and maintenance while in lytic, where most of EBV genes are expressed, the viral DNA is being replicated and packed in virions to produce new viral particles. EBV latency can be divided into four phases (latency 0, I, II and III) based on the set of genes expressed at each stage.

In contrast, EBV lytic induction initiates a rapid cascade of genes expression to facilitate the production of new viruses. Latency and lytic cycles will be discussed in more details in the following sections.

#### **1.3.4.2 EBV primary infection**

EBV can be detected in the saliva of healthy infected carriers and is typically transmitted through the saliva from one person to another. However, whether EBV infects first B cells or epithelial cells in the oropharynx *in vivo* is still unclear. It is widely documented that EBV exists as a cell-free virus as well as a cell-associated virus in the oropharynx (Haque and Crawford, 1997). Also, the presence of virus-producing epithelial cells *in vivo* is well-documented (Lemon et al., 1977, 1978; Sixbey et al., 1984). EBV particles found in the saliva have similar characteristics to viruses shed from epithelial cells which has higher viral envelope glycoproteins gp42 (Jiang et al., 2006). However, recent studies showed that cell-free EBV is less efficient at infecting epithelial cells and requires resting B cells to mediate this infection (Shannon-Lowe and Rowe, 2011; Shannon-Lowe et al., 2006). This adds ambiguity to the cells targeted in primary EBV infection.

The current model of EBV primary infection suggests that upon entry through saliva, infectious virus particles cross epithelial barriers to directly infect surrounding non-dividing naive B cells. This activates EBV-infected B cells into proliferating latent cells expressing all EBV latent genes. The expression of all EBV latent genes is known as type III latency program; latency will be discussed in the next sections.

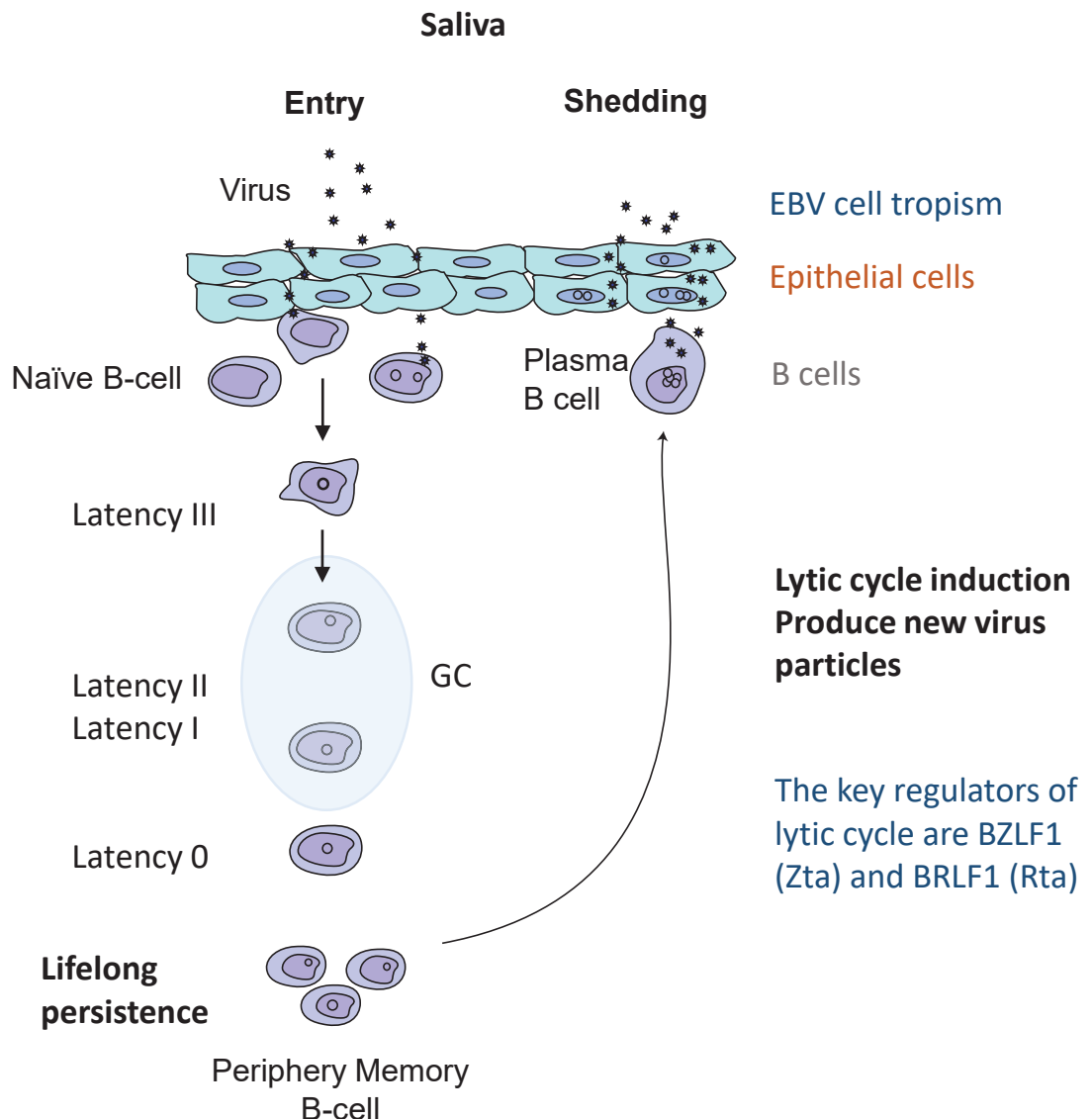
EBV newly infected B cells move into the germinal centre (GC) where they differentiate into memory B cells and EBV latent expression shifts to a more restricted latency expression profile. These cells join the memory B cell pool and establish EBV life-long persistence within the host (reviewed in (Chesnokova et al., 2015; Thorley-Lawson, 2015; Thorley-Lawson et al., 2013; Young and Rickinson, 2004)).

Upon antigen-activation of memory B cell infected cells, they become terminally differentiated B cells (plasma cells) capable of producing new virions that are either shed into saliva by B cells surrounding the epithelial oropharynx or transferred to polarized basal surface epithelia (Shannon-Lowe and Rowe, 2011; Shannon-Lowe et al., 2006). Moreover, it is suggested that EBV released from productive B cells within the oral lymphoid tissue infects other nearby cells such as monocytes which subsequently migrate into the oral epithelia and transfer the virus to the epithelial cells (Tugizov et al., 2007). Once the virus is transferred into the oral epithelium, viral replication and shedding into saliva begins in terminally-differentiated epithelial cells (Allday and Crawford, 1988).

The fact that epithelial cell terminal differentiation supports EBV replication comes from studies in oral hairy leukoplakia (OHL) (Greenspan et al., 1985). In OHL EBV DNA, lytic proteins, as well as latent proteins, were detected in the mid-upper layer of OHL lesions but not the basal levels (Löning et al., 1987; Niedobitek et al., 1991; Sandvej et al., 1992; Zhang et al., 1988). For example, BZLF1 protein, which is essential for lytic cycle induction, was only detected in the upper layers of OHL lesions but not the basal layers (Young et al., 1991b).

Despite the initial failure to detect EBV DNA using in situ hybridization in OHL basal layers (Niedobitek et al., 1991), it is believed that EBV infects and establishes latency in undifferentiated epithelia. Therefore, virus shedding only occurs at the mucosal surface layer (Allday and Crawford, 1988; Nawandar et al., 2015; Wille et al., 2013). **In Figure 1-3**, a model of EBV primary infection and biphasic life cycle is illustrated, showing the virus major events that lead to viral latency.





**Figure 1-3 EBV life cycle.** The main route of EBV transmission is through saliva. EBV infection in naïve B cells leads to their activation. The activated B cells migrate into the germinal centre (GC). Changes to EBV expression profile take place at this stage, and a few EBV-infected cells join the memory B cell pool. In memory B cells, EBV expression becomes restricted to allow for lifelong persistence (latency 0/I). Activation of EBV-infected memory B cells triggers their differentiation into plasma cells. Cellular factors associated with this differentiation activate BZLF1 (Zta) which leads to the productive EBV (lytic) cycle. Newly produced virions in the plasma cells infect oral epithelial cells. Differentiated oral epithelia allow EBV lytic activation thus acting as further amplifiers of EBV. New EBV virions are shed from oral epithelia into saliva.

#### 1.3.4.1 EBV epigenetic regulation

Like other herpesviruses, EBV DNA enters the nucleus as a naked linear DNA lacking any chromatin-associated proteins or DNA CpG methylation. This linear DNA becomes a covalently closed circular episome approximately 20 hours post-transfection. The rapid circularisation process helps to protect the linearised DNA ends from degradation and also reduces any DNA damage response in the host cell (Hurley and Thorley-Lawson, 1988). A delayed DNA damage response (DDR) up to seven days post-transfection is observed in EBV-infected cells. It is thought that this DDR is not related to EBV linear DNA but as a consequence of induced stress by cellular proliferation (Nikitin and Luftig, 2012; Nikitin et al., 2010).

The naked structure of DNA in the early infection stage (also known as pre-latent phase) is thought to allow transcription at viral promoters. Thus, soon after the entry of EBV DNA into the nucleus, several EBV proteins are expressed such as the latent proteins EBNA1, EBNA2, and EBNA-LP (Hammerschmidt, 2015). EBNA2 mRNA expression, one of the earliest expressed protein in EBV infection, is driven from the Wp promoter. EBNA-LP is co-expressed with EBNA2 from the same promoter. Transcription from this promoter is mainly regulated by the B cell factor PAX5. EBNA2 causes the switch from the Wp promoter to an alternate promoter (Cp) which leads to the expression of a longer transcript that includes EBNA1 and EBNA3 genes (EBNA 3A, 3B & 3C). EBNA1 can be expressed independently from a separate promoter (Qp) in the absence of other EBNAs. EBNA2 activates transcription of the viral latent membrane proteins LMP1 and LMP2 (reviewed in (Lieberman, 2015). The genomic locations of EBV latent promoters and transcripts are indicated in **Figure 1-2**. At this pre-latency stage, viral genes expression at this stage is not restricted to these latent proteins. Many lytic cycle genes are also transcribed. These include *BZLF1* and *BRLF1* (Kalla et al., 2010) as well as the early lytic genes such as *BHRF1*, *BALF1*, *BMRF1*, *BCRF1*, and *BNLF2a* (Hammerschmidt, 2015). The unmethylated state of EBV DNA at this stage inhibits the induction of EBV lytic cycle and DNA replication despite the expression of the *BZLF1* (Kalla et al., 2012; Kalla et al., 2010). Many of the

expressed proteins at this stage facilitate B cell proliferation and apoptosis escape (Woellmer and Hammerschmidt, 2013). For instance, BZLF1 and BNLF2a help to avoid apoptosis and host immune response, respectively (Hislop et al., 2007; Inman et al., 2001). As described earlier, some of these proteins, e.g. BZLF1 and BNLF2a, are found in the virion as packed RNA, further illustrating their importance at this stage (Jochum et al., 2012). The pre-latency phase takes around 10-14 days post-EBV infection (Hammerschmidt, 2015).

Epigenetic regulation of EBV DNA plays an essential role in EBV life cycle control bringing the EBV pre-latent phase to an end (reviewed in (Hammerschmidt, 2015; Lieberman, 2015)). The viral infection phase switches from pre-latent stage to latent stage. The EBV DNA becomes chromatinised and acquires DNA methylation. EBV DNA is chromatinised in a very similar way to the host genome; resembling a highly complex, and heterogeneous chromatin structure (Dyson and Farrell, 1985). Moreover, DNA CpG methylation is detected widely across the EBV episome (Diala and Hoffman, 1983; Dyson and Farrell, 1985). These changes in the EBV genome are also synchronised with changes within the EBV-infected cell itself. For instance, as shown in **Figure 1-3**, once EBV infects naïve B cells they rapidly become activated and enlarged, to begin proliferation.

Nucleosome occupancy of viral DNA starts very early during the pre-latency stage around four days post-transfection. CTCF binding sites across the EBV genome are believed to be key players in organising EBV chromatin structure (Arvey et al., 2012; Lieberman, 2015). On the contrary, DNA methylation of viral DNA is a slow process in infected primary B lymphocytes. CpG methylation becomes first detectable after two weeks and continues to increase for three months after EBV infection (Hammerschmidt, 2015; Kalla et al., 2010).

DNA methylation and histone remodelers are critical in controlling EBV life cycle. DNA methylation is involved in silencing most latent EBV promoters with the exception for a few active sites such as Q promoter (Hammerschmidt, 2015). Nucleosomes are dense on repressed lytic promoters during latency

ChIP analysis for histone 3 at various lytic promoters indicated a loss in nucleosomes occupancy at these promoters during lytic cycle activation (Woellmer et al., 2012). Furthermore, repressed lytic cycle promoters carry repressive chromatin marks such as H3K9me3 and H3K27me3. The latter, which is a polycomb repressor complex-associated mark, is reported to decrease significantly upon lytic cycle induction. ChIP analysis for EZH2, the H3K27me3 methyltransferase, validated the changes seen in H3K27me3 during EBV lytic cycle. In agreement with this, the chromatin activation marks such as histone acetylation and H3K4me3 are reported to increase at lytic promoters during lytic cycle induction (Murata et al., 2012; Woellmer et al., 2012). Histone repressive marks are also present on *BZLF1* promoter (Zp) such as H3K27me3, H3K9me2/3 and H4K20me3 in Raji cells (Murata et al., 2012) and H3K27me3 and H3K9me3 in Akata cells (Ramasubramanyan et al., 2012b).

In conclusion, EBV DNA chromatin structure governs the viral life cycle. In the early events of EBV infection, the key lytic cycle inducer BZLF1 is expressed. However, this does not lead to a viral productive life cycle because of the lack of viral DNA CpG methylation which is required for BZLF1 induced lytic cycle. The viral genome then becomes heavily associated with nucleosomes, CpG methylation, and chromatin repressive marks to facilitate a latent phase and establish persistence. During the productive phase, these marks have to be reversed to allow the transcription of EBV lytic genes that contributes to the production of new virus particles. This phase is regulated by BZLF1 which can bind CpG methylated motifs and recruit histone modifiers such as CBP (discussed further in **Section 1.4**).

#### 1.3.4.2 EBV latency and transformation

As discussed in earlier sections, latency is divided into different stages (Latency 0, I, II and III) based on the set of expressed genes. The full latency profile (latency III) includes the expression of a set of six polycistronic nuclear proteins known as EBV nuclear antigens (EBNA1, 2, 3A, 3B, 3C, and EBNA-LP), three proteins found in the cell membrane (LMP1, 2A, and 2B) and several small

RNAs, specifically, the non-polyadenylated EBV-encoded small RNAs (EBERs), the noncoding RNA transcripts of the BamHI rightward transcripts (BARTs) and multiple miRNAs (reviewed in (Longnecker et al., 2013; Young and Rickinson, 2004)) (see **Table 1.3** for different EBV latency stages).

Latency stage	Genes expressed	Associated cells and diseases
Latency 0	EBERs and miRNA/BART transcripts	Peripheral memory B cells and non-dividing Burkitt's lymphoma cells
Latency I	EBERs, miRNA/BART transcripts, and EBNA1	Dividing memory B cells and Burkitt's lymphoma
Latency II	EBERs, miRNA/BART transcripts EBNA1, LMP1 and LMP2A/B	Germinal centre B cells, Hodgkin lymphoma, nasopharyngeal carcinoma (NPC) and gastric carcinoma
Latency III	EBERs, BHRF1 miRNAs, miRNA/BART transcripts, EBNA1, EBNA-LP, EBNA2, EBNA3s, LMP1 and LMP2A/B	EBV-infected naïve B cells, post-transplant lymphoproliferative diseases (PTLD) and Lymphoblastoid cell lines (LCLs) ( <i>in vitro</i> )

**Table 1.3 Latency program.**

Data collected from (Longnecker et al., 2013; Price and Luftig, 2014).

The importance of latency is not only seen in the ability of the virus to hide and establish life-long persistence but in its ability to readily hijack various cellular pathways and immortalise infected cells. Thus, much of EBV research is focused on understanding the role of different latency proteins as these

contribute to our understanding of how EBV-associated cancers develop. The viral latency proteins have evolved to be specialised in mimicking cellular pathways, driving cell growth, transformation, chromatin remodelling, and immune evasion.

#### The role of EBV latent proteins

The role of EBV latency proteins in B cells transformation is evident from the ability of EBV to transform and establish latency III in B cells *in vitro* generating lymphoblastoid cell lines (LCLs) (Henle et al., 1967; Nilsson et al., 1971). Latency III EBV-infected B cells *in vivo* are associated with the development of lymphomas in immunocompromised patients. These cells, which are highly immunogenic, are readily cleared by the host immune system in healthy individuals. B cells expressing various latency programs *in vivo* have been detected in different disorders. Latency II has a more restricted expression profile limited to EBNA1, LMP1, LMP2 and the noncoding RNAs. This latency program is seen in NPC and HL (discussed in **Section 1.3.6**). In latency I, which is associated with BL, only EBNA1 and the noncoding RNAs are detected while the latter are detected exclusively in Latency 0. Latency program 0 and I do not lead to any immune response (Longnecker et al., 2013; Young and Rickinson, 2004). The role of each latent protein/RNA will be discussed briefly in the following paragraphs.

#### EBNA1

The 641 amino acids EBNA1 protein, expressed in all latency programs, has a primary role in mediating EBV genome replication and persistence in dividing cells (Yates et al., 1984). EBNA1 is transcribed from the Q promoter (Qp) in the restricted latency programs where the other EBNAs are not expressed. EBNA1 binds OriP and tether the EBV genome to cell chromosomes to mediate the replication of the EBV genome once per cell cycle. EBNA1 also acts as a transcription activator inducing the expression of latent genes from the viral C promoter (Cp). The Cp is also adjacent to OriP which contains multiple binding sites for EBNA1 (Frappier, 2015). EBNA1 plays a role in cancer pathogenesis as it disrupts promyelocytic leukaemia nuclear bodies (PML), decreases p53 activation, inhibits the NF-κB pathway, and enhances AP1 activity in NPC.

Unlike other EBNA proteins, EBNA1 peptides are not presented on the major histocompatibility complex class I (MHC I) so that it does not trigger any immune response by CD8 T cells. This is mainly attributed to the Gly-Ala repeats within EBNA1, which also explains the survival of cells displaying latency I *in vivo* (Frappier, 2015; Kang and Kieff, 2015).

### EBNA-LP

EBNA-LP is expressed from the W repeats (W1 and W2 exons) within IR1 as mentioned earlier. Expression of EBNA-LP can be driven from either the Wp (upon infection) or the C promoter in later EBV infection phases in LCLs. The switch in promoters is attributed to EBNA2 expression which regulates C promoter. EBNA-LP and EBNA2 are transcribed as bicistronic RNA. The EBNA-LP protein varies in size based on the number of W repeats (Allan and Rowe, 1989). The primary function of EBNA-LP is to co-activate transcription with EBNA2 at various promoters (Kang and Kieff, 2015; Kempkes and Ling, 2015).

### EBNA2

EBNA2 associates with cellular proteins such as RBPJ (also known as CBF1) and PU.1, to mediate transcriptional regulation through its two transactivation domains. EBNA2 stabilises RBPJ binding to DNA, recruits general transcription factors such as TFIIB, TAF40 as well as various coactivators and histone acetylases proteins such as CBP/P300 (Kempkes and Ling, 2015). The numerous targets of EBNA2 indicate the importance of EBNA2 in regulating viral and cellular genes. The Wp is the first active promoter upon EBV infection that leads to the expression of the EBNA proteins (Cheung and Kieff, 1982; Kempkes and Ling, 2015). EBNA2 is known to activate and mediate the switch to the C promoter. EBNA2 also activates the expression of LMP1 and LMP2 proteins. Furthermore, at the cellular level, many EBNA2 identified targets are critical for cell transformation, viability and proliferation such as NF- $\kappa$ B complex as well as *RUNX1*, *RUNX3*, *CXCR7* and *c-Myc genes* (Kempkes and Ling, 2015). In recent years, the global association between EBNA2 and the cellular genome has been addressed using genome-wide ChIP-Seq analysis. These experiments identified thousands of EBNA2 binding sites within proximal promoters and distal enhancers across the genome; many of these sites were

also shown to be targeted by EBNA3s (McClellan et al., 2013; Portal et al., 2013). More recently, EBNA2 was shown to associate with various super enhancer elements (Zhou et al., 2015) and regulate RUNX1 and RUNX3 through distal super enhancer elements (Gunnell et al., 2016).

The EBNA3 family comprises three proteins EBNA3A, EBNA3B and EBNA3C. These proteins bind to DNA and associate with RBPJ acting as a coactivator/repressor of EBNA2. Both EBNA3A and EBNA3C are essential for B cell transformation but EBNA3B is not. EBNA3A and EBNA3C promote cell proliferation by repressing the expression of pro-apoptotic proteins such as Bim (also known as BCL2L11) as well as tumour suppressors such as p16. In contrast to EBNA2, EBNA3A and EBNA3C inhibit RBPJ recruitments to DNA, thus, interfere with EBNA2 activation which is mediated through RBPJ (reviewed in (Allday et al., 2015; Kang and Kieff, 2015)).

### LMP1 and LMP2

LMP1 and LMP2 are two latent essential proteins for cell immortalisation. Latent membrane protein 1 is a primary EBV oncogene and has six transmembrane domains. These domains ensure stable association with the membrane glycolipoproteins and mediate constitutive cell signalling. LMP1 mimic surface CD40 on B cells, which mediate the activation of CD4 T cells (T helper cells). LMP1 expression is reported in HL, LCLs, undifferentiated NPC and during EBV lytic cycle (Kang and Kieff, 2015). LMP1 contributes directly to cell transformation; fibroblasts and B cells transfected with LMP1 show uncontrolled growth and proliferation (Kilger et al., 1998; Yang et al., 2000). It is known to activate various pathways such as the NF- $\kappa$ B and MAPK pathways. It also affects epithelial cell differentiation and epithelial-mesenchymal transition (EMT) in NPC cells (Kang and Kieff, 2015; Kieser and Sterz, 2015).

EBV encodes two additional similar latent membrane proteins, LMP2A and LMP2B. Both proteins have 12 transmembrane domains but LMP2B lacks a cytoplasmic signalling domain. These two proteins associate with the B cell Lyn and Syk protein tyrosine kinases. LMP2A blocks normal B cell receptor (BCR) signal transduction by mimicking antigen-independent activated BCR signalling. This induces B cell growth, transformation and promotes the survival of B cells



that are lacking BCR during germinal centre selection. In LCLs, LMP1 blocks the activation of tyrosine kinases, which inhibit EBV lytic cycle activation (Cen and Longnecker, 2015; Kang and Kieff, 2015). In epithelial cells, LMP2A and LMP2B together promote epithelial-mesenchymal transition through the activation of integrin and increase global tyrosine phosphorylation in these cells (Allen et al., 2005; Cen and Longnecker, 2015).

#### EBV-encoded small RNAs

In addition to the proteins above, EBV expresses two EBV-encoded small RNAs known as EBER1 and EBER2 (Lerner et al., 1981; Rymo, 1979). These non-polyadenylated non-coding RNAs are transcribed by host RNA polymerase III. EBERs are highly abundant and up to 5 million copies can be found per cell (Lerner et al., 1981). EBERs function remains obscure. However, they might play a critical role in B cell transformation. Viruses that harbour a deletion for EBERs lose 50% of their transforming capacity (Yajima et al., 2005). Also, expressing EBERs in EBV-negative Akata cells restores their tumorigenic ability in SCID mice (Komano et al., 1999).

EBERs bind various cellular proteins including the pattern-recognition receptors, autoantigen La, retinoic acid-inducible gene I (RIG-I), and AU-rich element binding factor 1 (AUF1) (Kang and Kieff, 2015; Lerner et al., 1981; Samanta et al., 2006; Skalsky and Cullen, 2015). EBERs are recognised by RIG-I, a cytosolic protein which senses viral dsRNA. RIG-I is part of the innate immune response to protect against viral infection by activating protective cellular genes such as type I interferons (INFs) (Samanta et al., 2006). RIG-I detection of EBERs also induces anti-inflammatory cytokines such as IL-10 which promotes cell growth. The viral-induced IFNs induce RNA-dependent protein kinase (PKR), which leads to the inhibition of the translation machinery. However, EBERs bind PKR and block its activation (Kang and Kieff, 2015).

#### EBV miRNAs

EBV expresses at least 25 pre-miRNAs that form approximately 40 mature miRNAs. EBV miRNAs are transcribed from within the BART locus introns (21 miRNAs) and from BHRF1 ORF (3 miRNAs) (Cai et al., 2006; Pfeffer et al., 2004; Skalsky and Cullen, 2015). The expression level of EBV miRNAs is

variable between different cells and latency stages. EBV BART miRNAs are expressed in all latency stages while BHRF1 miRNAs are more restricted to Latency III (Pratt et al., 2009; Yang et al., 2013). Most EBV miRNAs facilitate viral persistence and cell transformation in EBV-infected cells. It is worth noting that the majority of the BART region is deleted in EBV B95-8 isolate. However, this does not affect the transformation ability of this EBV strain. EBV miRNAs target many cellular mRNAs; for instance, the BART miRNAs were reported to target PUMA and Bim, two pro-apoptotic proteins. Furthermore, cells infected with EBV that harbour a deletion for BHRF1 miRNA grew slower compared to cells infected with wild-type EBV. This suggests a role for EBV miRNAs in cell transformation. BHRF1 miRNAs were reported to target cellular genes involved in triggering a host immune response such as CXCL-11, a T cell-attracting chemokine. The expression and function of EBV miRNAs have been extensively reviewed in (Kuzembayeva et al., 2014; Skalsky and Cullen, 2015).

A few EBV miRNAs were reported to target viral latent and lytic mRNAs. For instance, miR-BART20-5p was shown to directly target the immediate early lytic BZLF1 and BRLF1 mRNAs (Jung et al., 2014). Moreover, miR-BART2-5p was reported to target BALF5, the viral DNA polymerase (Barth et al., 2008; Pfeffer et al., 2004). EBV miRNAs targeting viral lytic genes aid in suppressing EBV lytic cycle and promoting latency.

### EBV BARTs

EBV is known to have several BamHI A rightward transcripts (BARTs) (also known as complementary strand transcripts (CSTs)). These are highly spliced polyadenylated mRNAs that span at least 16 overlapping exons (over 20 kbp) and arise from the same promoter (de Jesus et al., 2003; Sadler and Raab-Traub, 1995; Smith et al., 2000). These various mRNAs were first reported in NPC cells passaged in nude mice. Several ORFs were identified known as BARF0, RK-BARF0, RPMS1 and A73; however, results regarding the detection of an encoded protein for these mRNAs were contradicted (Raab-Traub, 2009; Smith, 2000). BART mRNAs are widely expressed in EBV-positive epithelial and B cell lines. Also, they are detectable by PCR in the peripheral blood of healthy EBV carriers (Chen et al., 2005; Chen et al., 1999).

### 1.3.4.3 EBV lytic cycle

Once EBV lytic cycle is activated, a cascade of about 80 viral genes are expressed in a temporally regulated manner (Kenney, 2007; Longnecker et al., 2013; Yuan et al., 2006). The first viral genes to be transcribed are the viral immediate-early (IE) genes *BZLF1* and *BRLF1* (Miller et al., 2007); both proteins are essential to complete EBV lytic cycle (Feederle et al., 2000). *BZLF1* and *BRLF1* expression in most latently infected cells induces EBV lytic cycle (Chevallier-Greco et al., 1986; Countryman and Miller, 1985; Darr et al., 2001; Ragoczy et al., 1998). The immediate-early lytic genes are characterised by the direct expression upon viral infection without being dependent on *de novo* viral or cellular protein synthesis. This was confirmed for *BZLF1* where transcription was reported to be as early as 1.5 hours post-EBV infection in EBV-negative BL cell lines and primary B cells in the presence of a protein synthesis inhibitor (Amon et al., 2004; Wen et al., 2007).

The viral IE proteins, either individually or cooperatively, activate the expression of the viral early (E) genes (Chevallier-Greco et al., 1986). For example, BMRF1, the viral DNA polymerase processivity factor, is activated more efficiently in the presence of both BZLF1 (Zta) and BRLF1 (Rta) (Cox et al., 1990; Holley-Guthrie et al., 1990). In general, most EBV early genes act collectively to ensure successful completion of viral lytic DNA replication (Longnecker et al., 2013). This is clearly seen in EBV DNA replication where several core replication proteins are assembled at OriLyt. These proteins include the immediate-early proteins (BZLF1 and BRLF1), EBV DNA polymerase (BALF5), DNA polymerase processivity factor (BMRF1), EBV single-stranded DNA binding protein (BALF2), helicase (BBLF4), primase (BSLF1), and primase-associated factor (BBLF2/3) (El-Guindy et al., 2013; Fixman et al., 1992, 1995; Liao et al., 2005). Also, viral early lytic genes encode proteins such as viral kinase (BGLF4) (Chen et al., 2000), anti-apoptotic host BCL2 homologue (BHRF1) (Henderson et al., 1993), transcription factor (BRRF1) (Hong et al., 2004) and immune evasion protein (BNLF2a) (Hislop et al., 2007). The early lytic genes are dependent on *de novo* protein synthesis. However, unlike EBV lytic latent genes, they are not dependent on viral DNA replication (Amon et al., 2004).

The late lytic viral genes are more associated with EBV virion structural proteins. Transcription of most late lytic genes is dependent on both *de novo* protein synthesis and viral DNA replication (Amon et al., 2004; Kenney, 2007). An exception to this is the transcription of BcLF1 and BFRF3 genes which are not dependent on viral DNA replication (Serio et al., 1997). There are approximately 36 late lytic genes including tegument proteins, glycoproteins and a viral IL-10 homologue (McKenzie and El-Guindy, 2015; Yuan et al., 2006). The late lytic promoters are characterised by the presence of non-canonical TATA box (TATT) and thought to be regulated differently from early lytic promoters (Serio et al., 1998). The expression of the late lytic genes requires the assembly of several viral early proteins including BcRF1 which is a viral TBP-like protein that binds the TATT motif (Aubry et al., 2014; Buisson et al., 1989; Gruffat et al., 2012; Gruffat et al., 2016; Serio et al., 1998).

Lytic cycle reactivation is associated with the differentiation of cells *in vivo*. In B cells, the activation of memory B cells into plasma cells through the B cell receptor induces downstream B cell factors that activate the expression of the viral immediate early proteins (Crawford and Ando, 1986; Laichalk and Thorley-Lawson, 2005). The human proteins XBP-1 and BLIMP-1, which are expressed during plasma cell differentiation, are thought to play a significant role in activating BZLF1 and BRLF1 promoters (Zp and Rp, respectively) (Bhende et al., 2007; McDonald et al., 2010; Reusch et al., 2015; Shaffer et al., 2002). Similarly, in epithelial cells, induced lytic cycle is linked with terminal differentiation (Karimi et al., 1995; Young et al., 1991a). In epithelial cells, the cellular factors BLIMP-1 and KLF4 synergise in inducing lytic cycle (Nawandar et al., 2015). The epithelial differentiation-dependent kruppel like factor 4 (KLF4) is capable of inducing lytic cycle by binding consensus motifs within both Zp and Rp promoters in epithelial as well as B cells (Murata et al., 2013; Nawandar et al., 2015).

*In vitro*, the transient expression of BZLF1 and BRLF1 can disrupt latency and induce EBV lytic cycle. In addition, several agents can activate lytic cycle such as the histone deacetylase (HDAC) inhibitor sodium butyrate, DNA methyltransferase inhibitors, and phorbol ester (McKenzie and El-Guindy, 2015). Also, in EBV-positive BL Akata cells, cross-linking BCR with anti-

immunoglobulin activates EBV lytic cycle (Takada and Ono, 1989). On the contrary, some compounds that specifically interfere with the viral DNA polymerase can inhibit EBV lytic cycle such as acyclovir and ganciclovir (Datta et al., 1980; Meng et al., 2010).

Understanding the lytic cycle of the virus is highly significant because this phase leads to the production of infectious virus particles, thereby, promoting the infection of new cells as well as the transmission of the virus to new hosts. Also, various lytic proteins have a diverse role in manipulating different cellular pathways. More importantly, high EBV load in peripheral blood and antibody titre of viral capsid antigen (VCA) are two features that are always associated with either the pre-onset or pathogenesis of EBV-associated diseases as will be seen in **Section 1.3.6**. Moreover, expression of the viral lytic cycle genes *BZLF1*, *BALF2*, and *BCLF1* were detected in NPC and healthy control biopsies while *BRLF1* was identified specifically in NPC patients (Feng et al., 2000). Similarly, *BZLF1* was detected in a subset of BL tumours and high anti-*BZLF1* IgG was also associated with BL (Asito et al., 2010; Niedobitek et al., 1995). These findings suggest a possible role for these lytic cycle genes in the pathogenesis of these diseases.

### **1.3.5 EBV cell tropism**

As mentioned earlier, EBV infects B cells as well as epithelial cells. However, there are many differences between the events that associate with EBV infection in both cells. EBV infection of B cells is better understood because B cells are easily collected from peripheral blood samples and EBV infects these cells very efficiently *in vitro* (Moore et al., 1989). In contrast, EBV is less efficient at infecting epithelial cells as they do not express the CD21 receptor (Cohen, 2000; Feederle et al., 2007).

The dynamic entry of EBV into B cells is facilitated through high-affinity interaction between the EBV virion glycoprotein gp350 and CD21 (also known as complement receptor type 2 (CR2)) on B cell surface (Fingerroth et al., 1984; Nemerow et al., 1987). Another EBV glycoprotein gp220, a splice variant of gp350 that also has a binding site to CD21, might aid in the attachment

process. CD21 also exists as a complex with surface receptors CD19 and CD81 on primary B cells. As the virus attaches, the crosslinking of CD19 and CD21 is mediated through gp350. This can activate the NF- $\kappa$ B and protein kinase C pathways which affect downstream events within the infected cells (Sinclair and Farrell, 1995). Also, EBV gp350 is known to interact with CD35, a complement receptor type 1, which helps in initiating EBV infection (Chesnokova et al., 2015).

Once attached to the B cell surface via CD21, the virus is endocytosed to form a membrane vesicle from which it escapes by fusing its envelope with the cell membrane. This fusion is mediated through the interaction of four EBV glycoproteins (gH, gL, gp42 and gB) with the cellular protein, HLA class II (Chesnokova et al., 2015). EBV gH and gL are two glycoproteins that dimerise together to give a correct folding of the final protein structure (Matsuura et al., 2010). Thus, they are commonly referred to as gHgL. The cleaved form of glycoprotein gp42 on EBV virion associates with gHgL and interacts with HLA II, which is essential for successful fusion (reviewed in (Chesnokova et al., 2015)).

In contrast to B cells, epithelial cells lack or express low levels of CD21 (Fingerroth et al., 1999), which mediates EBV attachment through viral gp350 in B cells as discussed above. The high affinity of CD21 to gp350 plays a critical role in EBV infection. Epithelial cell lines that were engineered to express CD21, showed improved EBV infection (Li et al., 1992). Furthermore, epithelial cells lack CD19 and CD35 which also play a role in EBV attachment in B cells.

In epithelial cells, EBV is believed to use gHgL to bind to a subset of integrin proteins that comprise the integrin  $\alpha$ -v and beta integrins. Additionally, EBV *BMRF2* encodes a glycoprotein that was also shown to interact with certain integrin proteins (Chesnokova and Hutt-Fletcher, 2014; Chesnokova et al., 2015).

EBV infection of epithelial cells *in vivo* is believed to be mediated by virus particles attached to CD21 on B cells interacting with epithelial cells at the basolateral surface. *In vitro* analysis showed epithelial cells are infected more

efficiently through B cell transfer infection. This interaction took place through the binding of the heparan sulphate moieties of CD44v3 to the lymphocyte-endothelial-epithelial cell adhesion molecule (LEEP-CAM) (Shannon-Lowe et al., 2009; Shannon-Lowe and Rowe, 2011).

Unlike in B cells, internalisation of the virus in epithelial cells is less understood due to the lack of an appropriate epithelial system. EBV gB, gH, and gL are required for epithelial fusion but not gp42. HLA II is not expressed in epithelial cells. Thus, the fusion is promoted by the interaction of gHgL and integrin proteins (Chesnokova and Hutt-Fletcher, 2014; Chesnokova et al., 2015). It is worth noting that EBV virions produced in epithelial cells are more enriched for gp42 making the virions more competent at infecting B cells but not epithelial cells. Similarly, virions produced in B cells are more efficient in infecting epithelial cells due to the reduced gp42 expression (Jiang et al., 2006).

### **1.3.6 EBV-associated diseases**

EBV infects and establishes life-long persistence in over 90% of the world population (Borer et al., 1999; de-The et al., 1975). EBV has been associated not only with the development of many cancers but also with several benign conditions such as infectious mononucleosis (IM) and oral hairy leukoplakia (OHL). EBV is recognised as the highly causative agent of lymphomas in immunocompromised patients and individuals with predisposing factors. EBV has been strongly associated with a wide variety of diseases. These diseases include Burkitt's lymphoma, Hodgkin's lymphoma (HL), nasopharyngeal carcinoma, T-cell malignancies, post-transplant lymphoproliferative disorder, gastric carcinoma (GC), breast cancer and several autoimmune disorders (reviewed in (Carbone et al., 2008; Chen, 2011; Kutok and Wang, 2006; Young and Rickinson, 2004)).

#### **1.3.6.1 Infectious mononucleosis (IM)**

Infectious mononucleosis (IM), also known as glandular fever, mono and kissing disease, is an acute disease caused by primary EBV infection in young adults



between the age of 15 -25 years (Dunmire et al., 2015). Most studies carried out in children at the onset of EBV infection have failed to report any major or similar symptoms to IM (Fleisher et al., 1979). A recent study, in which over 500 undergraduate students were monitored for EBV infection, showed that around 75% of the students that encountered EBV between the age of 18-22 have developed typical IM symptoms while about 15% developed atypical symptoms and 10% were asymptomatic (Balfour et al., 2013; Dunmire et al., 2015). Typical IM infection is associated with various symptoms including sore throat, general fatigue, fever, swollen lymph nodes and an acute increase in white blood cells count (Hickey and Strasburger, 1997).

The term 'Infectious mononucleosis' was suggested by Sprunt and Evans to describe the symptoms above over a 100 year ago (Butka, 1926). However, it is the same disease with glandular fever described by Emil Pfeiffer in 1889 (Tidy, 1950). The first indication that IM was associated with EBV occurred when a member of the Henle group, who used to give her blood as a negative control for EBV, had become seropositive for EBV after developing IM symptoms (Crawford, 2014). Further studies testing for EBV heterophile antibodies on various age groups over a long period showed that high antibody titre was associated with the conversion of seronegative to seropositive in IM patients (Henle et al., 1968). Although IM is mainly caused by EBV, it can also be associated with cytomegalovirus (CMV) (Hickey and Strasburger, 1997).

EBV primary infection, primarily transmitted through kissing in young adults (Balfour et al., 2013), is followed by an incubation period of about six weeks before the development of any symptoms. IM symptoms are caused by a remarkable expansion in EBV-specific CD8 T cells. IM is also characterised by a high EBV load in the oral cavity and peripheral blood accompanied by a sharp rise in IgM antibodies against EBV viral capsid antigen (VCA) (Balfour et al., 2013). EBV-infected B cells detected in tonsils in IM patients vary vastly in morphology and EBV latency gene expression (latency I-III). These cells include small to medium lymphoblasts as well as Reed-Sternberg cells similar to Hodgkin lymphoma (Kurth et al., 2000).



In general, the condition is self-limiting. However, in very rare cases complications such as haemolytic anaemia and thrombocytopenia were reported (Dunmire et al., 2015). There is no specific treatment for IM but corticosteroids can be prescribed to ease the inflammatory symptoms that accompany IM (Luzuriaga and Sullivan, 2010)

### **1.3.6.2 Burkitt's lymphoma (BL)**

Burkitt's lymphoma (BL) is categorised into three types: endemic BL, sporadic BL and immunodeficiency-related BL (Schmitz et al., 2014). As described in **Section 1.1**, endemic BL was first described by Denis Burkitt in a restricted geographical area overlapping the malaria region in Africa (Burkitt, 1958; Burkitt and Wright, 1963; Dalldorf et al., 1964). Endemic BL is strongly associated with EBV and commonly seen in children between 2-15 years of age. Typically, patients present with a large tumour growth of the jaw. However, tumours are also frequently detected in the abdominal area, in organs such as the liver and intestine (Burkitt, 1958; Burkitt and O'Connor, 1961). The association of EBV with endemic BL is over 98% in contrast to sporadic EBV-associated BL, which is only about 20% (Rochford and Moormann, 2015). Despite this strong association with EBV, BL aetiology is a debatable subject as the malaria incidence rate has also been strongly associated with this tumour (Dalldorf et al., 1964; Rainey et al., 2007).

A chromosomal translocation t(8:14) is recognised as a unique and driving factor of all types of BL (Hummel et al., 2006). This translocation places the oncogenic *c-Myc* gene under the control of the immunoglobulin (IG) heavy-chain locus (Dalla-Favera et al., 1982; Zech et al., 1976). This results in high constitutive expression of *c-Myc* through the IG enhancer elements. The role of *c-Myc* in BL pathogenesis was confirmed, using mouse models which showed that constitutive expression of *c-Myc* was capable of driving tumours in different B cell lineages (Adams et al., 1985; Sander et al., 2012).

EBV and malaria are the major risk factors in BL with a strong correlation between malaria and EBV infections in infant cases of endemic BL (Piriou et al., 2012). Nevertheless, the role of both infections in the development of BL

remains ambiguous. However, both diseases are common in causing B cell hyperplasia. Thus, as a consequence of increasing the pool of B cells, the chance of chromosomal translocations increases. This could, to some extent, explain the possible mechanisms in BL pathogenesis, although, other factors may also play a role (Magrath, 2012).

Multiple copies of EBV DNA were shown to be present in BL biopsies as well as cell lines derived from BLs (zur Hausen, 1972; zur Hausen et al., 1970). Additionally, BL is believed to be a result of an expansion of one single cell infected with EBV. This was observed upon the examination of EBV genome terminal ends patterns in more than 10 cases of endemic BL (Neri et al., 1991).

Also, it is well-established that EBV transforms B cells *in vitro* which results in immortalised LCLs (Henle et al., 1967). However, unlike BL, these LCLs express program latency III which includes nine EBV proteins as previously discussed. These proteins are responsible for B cell immortalisation (Bornkamm and Hammerschmidt, 2001). In endemic BL tumours, EBV remains latent in these cells expressing only the virus nuclear protein EBNA1 and several small untranslated RNAs such as EBERs and miRNAs from the BHRF1 and BART regions (latency I) (Bornkamm and Hammerschmidt, 2001). EBNA1 is mainly associated with EBV episome maintenance. However, it could also play a role in inducing lymphomas as demonstrated in a transgenic mouse model (Wilson et al., 1996). Additionally, EBNA1 might aid in B cell survival as it is widely undetectable by the host immune response. In contrast, EBERs were reported to play a role in increasing cells tumorigenesis and resistance to apoptosis (reviewed in (Brady et al., 2007; Magrath, 2012; Rochford and Moormann, 2015)). It is worth noting that an atypical latency profile was reported in some BL cases where the cells contained EBV genome with a deletion in the EBNA2 gene. These cells express EBNA1, EBNA3 and the viral Bcl2 homologue which aid in preventing apoptosis (Kelly et al., 2005).

Malaria has been linked with B cell hyperplasia as well as causing B cell activation through the parasite *Plasmodium falciparum* encoded protein CIDR1 $\alpha$ . This protein is expressed on the surface of infected red blood cells

leading to B cell activation and differentiation, thereby activating the EBV lytic cycle. This might play a role in increasing the viral load of EBV in endemic BL cases. Additionally, malaria could play a role in increasing the chances of chromosomal translocations through the interaction with Toll-like receptors (Magrath, 2012).

BL is treated with chemotherapy with a success rate of over 80% and very low mortality rate especially among young adults (Burkitt, 1965; Burkitt et al., 1965; Costa et al., 2013).

### **1.3.6.3 Hodgkin lymphoma**

Hodgkin's lymphoma (HL) is one of the most prevalent lymphomas in western countries with an annual rate of three cases per 100,000 (Murray and Bell, 2015). HL was first described by Sir Thomas Hodgkin in 1832, a lecturer in morbid anatomy at Guy's Hospital Medical School in London (Hodgkin, 1832; Stone, 2005). HL is classified into two types: classical (cHL), which accounts for around 95% of cases and nodular-lymphocyte predominant Hodgkin lymphoma (NLPHL) which forms only 5% of HL cases. The difference between the two types is based on the morphology of the tumor cells present in each type. In cHL, tumor cells are described as multinuclear Reed/Sternberg cells (HRS). These cells originate from pre-apoptotic germinal centre B cells and are characterised by a disrupted structure of lymph nodes, where malignant cells are surrounded by infiltrating cells such as T and B lymphocytes (Kuppers, 2012). cHL is subdivided into four morphological types: mixed cellularity, nodular sclerosis, lymphocyte rich and lymphocyte depletion (Murray and Bell, 2015).

EBV is mostly associated with cHL in which the virus DNA is detected in around 20-60% of cHL cases worldwide. Only a minority of NLPHL are believed to be associated with EBV (Kapatai and Murray, 2007). The association between EBV and cHL is widely variable between countries, gender and various age groups. For example, the percentage of EBV-positive cHL is about 57% in China, while lower rates have been reported in North America and Europe

(Kapatai and Murray, 2007). A distinctive mark of HRS cells in cHL is the lack of a functional B-cell receptor (BCR), which is a result of mutations affecting the immunoglobulin gene. Furthermore, these cells have defects in the expression of B cell-specific transcription factors such as PAX5, and might express markers of other cell types such as T cells. HRS cells are also characterised by high expression of NF- $\kappa$ B transcription factors, which prevents apoptosis and contribute to the tumorigenesis of these cells (Murray and Bell, 2015). These cells are also characterised by the consistent expression of various lymphocytes markers such as CD30, CD15, CD20, ID2 and Notch1 (Kuppers et al., 2012).

Similar to BL, monoclonal EBV genomes are detected in HL, which strongly suggests that HL results from the expansion of one cell that has been infected with EBV prior to tumorigenesis (Gulley et al., 1994). High EBV antibody titre and symptomatic IM are considered as risk factors in cHL development (Massini et al., 2009). It is not clear what role EBV plays in the development of HL, although, EBV may play a major role in promoting the survival of HRS cells. These cells express LMP1, LMP2A and -2B, EBNA1 and the EBER and BART transcripts (Latency II program). Although the HRS cells deficient BCR should promote apoptosis, LMP 1 and 2 proteins mimic normal CD40 and BCR signalling, respectively, contributing to the high expression of NF- $\kappa$ B and enhanced survival of these cells due to the activation of different pathways including RAS-PI3k-AKT pathway. Furthermore, the activation of EBV lytic cycle is prevented due to the lack of functional BCR signalling which is the major pathway that activates EBV lytic cycle and is required for B cell activation and differentiation to plasma cells (Murray and Bell, 2015).

HL cases are widely treated with either chemotherapy or radiation therapy with a good prognosis; 65%–90% of patients in the early stages of HL are expected to make a full recovery within five years of the treatment (Kuppers et al., 2012).

#### **1.3.6.4 EBV lymphoproliferative diseases (LPD) in immunocompromised individuals**

Patients with congenital or acquired immunodeficiency as well as post-transplant patients who are undertaking immunosuppressive treatments are at risk of developing EBV-related lymphomas. This is due to the uncontrolled outgrowth of EBV-infected B cells. In immuno-competent individuals, cells expressing EBV latency III program are immediately cleared by host immune response mediated by EBV-specific cytotoxic T cell, (Hopwood and Crawford, 2000).

A good example of such EBV lymphoproliferative disease is post-transplant lymphoproliferative disorder (PTLD) where patients are usually the recipient of an organ or stem cells transplant. PTLD is dependent on the type of transplant and the course of immunosuppression treatment. High incidence rates are common with heart, lung, small intestine or multiple organs transplant (Ok et al., 2015).

PTLD can range from either benign (IM-like) or non-Hodgkin malignant polyclonal lymphocytic proliferation. Morphologically, PTLD can be classified into early lesions, polymorphic PTLD, and monomorphic PTLD. Early lesions are further categorised into plasmacytic hyperplasia (increased plasma cell numbers) and IM-like lesions (similar to IM-B cell hyperplasia). On the other hand, polymorphic PTLD is characterised by lesions that affect the lymph node architecture. This is the result of an expansion of a polymorphic population of different lymphocytes including plasma cells and different B cells at various maturation stages (Ok et al., 2015). Unlike the polymorphic PTLD, monomorphic PTLD includes Burkitt's lymphoma, diffuse large B-cell lymphomas and classic Hodgkin lymphomas that developed in post-transplant patients (Carbone et al., 2008).

The pathogenesis of PLTD is linked with both EBV and KSHV (HHV-8). However, EBV is believed to be the major risk factor in PTLD as it is present in 60%-80% of PTLD. In common with other EBV lymphomas, in monomorphic PTLD, the EBV genome is monoclonal suggesting that EBV infection happens

before the expansion of the malignant cells (Carbone et al., 2008). High EBV viral load and reduced EBV-specific T cells are used as a predictive marker of PTLD (Smets et al., 2002). In PTLD, the EBV latency expression profile depends on the type of lymphoma, however, typically type III latency is detected. EBV infection contributes directly to the development of PTLD. This is due to the presence of cells that express the full set of EBV latent proteins EBNA, LMPs and BARTs RNA which are responsible for cell transformation and immortalisation.

Additionally, the EBV replicative cycle might play a role in PTLD as over 80% of tumours associated with immunosuppression treatment show evidence of EBV DNA replication (Montone et al., 1996; Tanner and Alfieri, 2001). Immunohistochemical staining of PTLD biopsies confirmed the presence of the immediate-early lytic protein BZLF1 (Zta) (present in 81% of specimens) and late lytic gp350 antigens, confirming a complete lytic cycle in PTLD (Montone et al., 1996).

Various treatment approaches are used in PTLD therapy including chemotherapy, and EBV viral load control. EBV viral load play an essential role in monitoring the disease and determine the best treatment approach. The overall survival rate of PTLD is estimated to be from 40%-60% within five years (Dierickx et al., 2015).

#### **1.3.6.5 Nasopharyngeal carcinoma**

Nasopharyngeal carcinoma (NPC) is a head and neck tumour and the most common cancer originating from the nasopharynx. NPC is rare worldwide but is prevalent in certain regions and among certain ethnic groups such as in Southern China, and North Africa (Chang and Adami, 2006). According to estimates from 2010, the incidence and mortality rates in Southern China were 19.5 and 7.7 per 100,000 persons, respectively (Yang et al., 2016).

NPC is classified into three subtypes: keratinizing squamous cell carcinoma (type I), non-keratinizing squamous cell carcinoma (types II) and

undifferentiated non-keratinizing squamous cell carcinoma (type III) (Young and Dawson, 2014).

The etiology of NPC is associated with many factors including EBV infection, host genetics, and environmental factors. Epidemiological studies tracing NPC cases among Chinese patients living in Hong Kong and Chinese immigrants in the USA suggest that diet is also a possible etiologic factor. In some areas in China, preserved food such as salted fish is commonly consumed. This type of food contains carcinogenic substances such as nitrosamines (Yu et al., 1981). This is further supported by the fact that such preserved food is also consumed in areas with high NPC incidence rates such as North Africa and Greenland (Raab-Traub, 2015). However, the fact that NPC is still high in Chinese immigrants born in the USA suggests that genetic factors also play a role in NPC (Buell, 1974). Indeed, various human leukocyte antigen (HLA)-DRB1 polymorphisms have been strongly linked with the etiology of NPC (Yang et al., 2016).

The association between EBV and NPC was first based on the specific high antibody titre for EBV seen in approximately 84% of NPC cases from Africa, China and USA (Henle et al., 1970; Old et al., 1966). A few years later, the EBV genome was detected in NPC cells (Wolf et al., 1975). EBV is present in all types of NPC but it is commonly associated with the differentiated form (Pathmanathan et al., 1995; Sixbey et al., 1984; Young and Dawson, 2014). Similar to BL and cHL, Southern blot analysis of EBV DNA terminal repeats in NPC cells confirmed the presence of monoclonal EBV genome. This indicates that EBV infection took place before the expansion of the first tumor cell (Raab-Traub and Flynn, 1986).

EBV remains latent (type II latency) in NPC cells expressing only EBNA1, BART RNAs, EBERs and the latent membrane proteins, LMP1 and LMP2 (Young and Dawson, 2014). NPC cells are characterised by the reduced expression of typical epithelial factors such as catenin and N-cadherin, which induces epithelial to mesenchymal transition (EMT) indicating the metastasis capacity of this cancer. EBV proteins and miRNAs aid in altering cell growth regulation. For



example, in NPC cells, EBNA1 might play a role in increasing genetic instability by disrupting promyelocytic nuclear bodies (PML), which aid in DNA damage repair. LMP1 is known to block p53-induced apoptosis while LMP2 might play a role in inducing epithelial cells transformation and inhibiting differentiation. Additionally, LMP2 induces cell migration by interfering with integrin expression and localization while BART miRNAs target pro-apoptotic cellular genes such as Bim (Raab-Traub, 2015).

EBV strain variations isolated from various NPC cases may also play a role in NPC pathogenesis. For instance, an EBV strain isolated in Chinese NPC biopsies had a variant LMP1 harbouring a 10 amino acid deletion (Young and Dawson, 2014).

EBV is an important tool in NPC diagnoses and prognosis. NPC is characterised by the presence of high IgA antibodies against EBV viral capsid antigen (VCA), and early antigen (EA). Also, EBV plasma viral load has been suggested as a useful marker in assessing NPC tumor load and stage (Chen et al., 2015). More recently, antibodies against the viral glycoproteins gH, and gL, which are required for EBV-epithelial cell infection, showed promising sensitivity in diagnosing NPC (Li et al., 2016). A combination of chemotherapy and radiotherapy is the current standard method for treating NPC. The overall five-year survival rate for early NPC stage patients is over 90%. However, this rate decreases with NPC late stages (Zhang et al., 2013).

#### **1.3.6.6 Oral hairy leukoplakia (OHL)**

Oral hairy leukoplakia (OHL), also referred to as hairy leukoplakia (HLP), is a benign condition associated with epithelial cells of the tongue and is mostly seen in HIV-infected immunocompromised patients. Interestingly, EBV is not detected in tongue tissue of immunocompetent individuals (Pagano, 2009). Patients present with white lesions on their tongue that result from the high EBV replication level in these lesions (Greenspan et al., 1985). Early studies on the role of EBV in OHL have helped to broadening our understanding of the EBV infection cycle in epithelial cells. BZLF1 protein was detected in the upper layer in OHL biopsies but not in lower layers indicating a link between BZLF1



expression and the degree of epithelial cells differentiation (Young et al., 1991c). These initial findings were further supported by the activation of *BZLF1* promoter (Zp) by inducing differentiation in squamous epithelial cells (SCC12F) (Karimi et al., 1995). Moreover, many other EBV transcripts were found to be abundant in OHL biopsies including *BCRF1*, *BKRF4*, *BMRF2* and *gp350/220* (Lagenaur and Palefsky, 1999; Lau et al., 1993). Further studies on *BMRF2* promoter showed high basal level expression of this promoter in epithelial cells which explains its high abundance in OHL (Lagenaur and Palefsky, 1999). In addition to lytic genes expression, the latent membrane protein LMP1 is also detected in OHL. LMP1 could play a role in the development of this disease by activating many proteins such as NF- $\kappa$ B, epidermal growth factor receptor (EGFR) and c-Jun terminal kinase 1 (JNK1) through the TRAF signalling pathway (Webster-Cyriaque et al., 2000).

OHL is resolved quickly within 1-2 weeks of treatment with any EBV viral replication inhibitors such as acyclovir. However, recurrence of the disease after successful treatment is commonly reported (Resnick et al., 1988; Walling et al., 2003).

#### **1.3.6.7 Gastric carcinoma (GC)**

Evidence that EBV was associated with gastric carcinoma (GC) was first obtained by the examination of around one thousand GC biopsies. 7% of the examined GC biopsies showed the expression of EBER1, and the EBV genome was detected in each cancer cell but not in surrounding lymphoid or mucosa stromal cells (Imai, 1994). This strongly suggested a role for EBV in EBV-positive GCs. In common with other EBV lymphomas, EBV genome terminal repeat analysis indicated a monoclonal origin for these cancers (Imai, 1994). The EBV genes expressed in gastric carcinoma are (EBNA1), EBER1 and EBER2, and the BARTs transcripts; a small amount of latent membrane protein 2A (LMP2A) is also expressed in some cases (Takada, 2000). EBV-positive GC patients had high serum EBV-specific antibodies, but EBV-specific cytotoxic T-cells were not detected. This strongly correlates with the latency expression profile in which EBNA1 is known to evade any host immune response (Imai,

1994). Unlike Burkitt's lymphoma and NPC, EBV positive gastric carcinoma cases are seen worldwide with no definite endemic regions; the incidence rate varies among regions ranging from 16–18% in the USA to 4.3% in China (Takada, 2000). Interestingly, *Helicobacter pylori* infection and smoking are recognised as the main risk factors for gastric carcinoma. Unlike NPC, there are no genetic predisposing factors identified in gastric carcinoma (Takada, 2000).

Recently, it has been suggested that EBV-associated gastric carcinoma (EBVaGC) should be recognised as a separate subtype of gastric carcinomas (GCs) based on the unique molecular presentation of these GCs. The molecular markers of EBVaGC include recurrent mutations in the kinase gene PIK3CA, and DNA hypermethylation (Cho et al., 2016). Morphologically, EBVaGC is characterised by marked dense lymphocytic infiltration in which the number of tumour-infiltrating lymphocytes is greater than that of tumour cells. EBVaGC is subdivided into three types according to the cellular immune responses: Lymphoepithelioma-like carcinoma (LELC), carcinoma with Crohn's disease-like lymphoid reaction (CLR), and conventional adenocarcinoma (CA) (Cho et al., 2016).

Currently, there is no specific treatment for EBVaGC. However, the primary option is surgical removal of the tumour. The use of EBV lytic cycle activation to induce cell lysis has been proposed. DNA demethylation agents can be used to target the hypermethylated EBV genome in EBVaGC; thereby inducing EBV lytic cycle (Iizasa et al., 2012).

#### **1.3.6.8 Other EBV-related diseases**

EBV may also play a role in a broad spectrum of benign, malignant, and autoimmune diseases. Here, some of these disorders will be discussed briefly.

Firstly, although EBV infects and establishes latency in B cells, EBV infection in other lymphocytes, such as T cells and natural killer (NK) cells, has been reported. In fact, EBV-associated T cell lymphoproliferative diseases have been known for over 25 years. Patients with such diseases were reported to have a substantial increase in EBV viral capsid antigen (VCA) antibody titre (Jones et

al., 1988). Unlike EBV-associated B cell lymphomas, T cell lymphomas that are associated with EBV are very rare and show site-dependent patterns. EBV is mostly prominent in virus-associated hemophagocytic syndrome (VAHS) T cell lymphoma and nasal T/NK lymphomas. Similar to other EBV-associated lymphomas, VAHS T cell lymphoma are infected with monoclonal EBV and express the latency I/II set of genes. VAHS T cell lymphoma might result in the life-threatening fatal complication of hemophagocytosis (Kutok and Wang, 2006), in which hyperinflammatory cytotoxic T cells and NK cells mediate the engulfment of various hematopoietic cells by activated macrophages (George, 2014).

Similarly, nasal NK/T cell lymphomas have clonal EBV genome in all cases. These aggressive lymphomas are very rare throughout the world but commonly seen in Southeast Asian and South American populations (Kutok and Wang, 2006; Tse and Kwong, 2013). Cells in these lymphomas, in most cases, are characterised by the lack of surface CD3 and T cell receptor as well as the presence of the NK marker CD56 and the expression of perforin, granzyme B, and TIA-1 (Kutok and Wang, 2006; Tse and Kwong, 2013). This correlates mostly with a NK cell origin for these tumours. However, in a small number of cases, EBV-associated T/NK lymphoma cells expressing CD2, CD3, and lack CD56 have been reported (Kutok and Wang, 2006; Tse and Kwong, 2013).

The role of EBV in T and NK cells oncogenesis is not fully understood and there is no *in vitro* model for EBV immortalisation of such cells. However, various EBV latent genes might play a role in transformation. For example, LMP1 interacts with TRAF pathway proteins to activate and upregulate several proteins including NF- $\kappa$ B and BCL2 which leads to cell survival. EBER and BART RNAs induce various growth cytokines and promote cell proliferations (Cai et al., 2015). It is worth noting that plasma EBV DNA load is correlated with tumour load, and high EBV DNA correlates with poor prognosis. These tumours are treated with a combination of chemo- and radiotherapy (Tse and Kwong, 2013).

An association of EBV infection with breast cancer remains to be proven despite many reports showing the presence of EBV DNA or EBERs in breast

cancer tissues. Contradicting results are presented in the literature regarding the correlation of EBV with breast cancer. For example, a recent meta-analysis of published reports found that 29.32% of breast cancer samples were infected with the Epstein-Barr virus and that EBV was more prominent in Asia (35.25%) compared to USA samples (18.27%) (Huo et al., 2012). However, a more recent study analysed EBV DNA in peripheral blood and tumour biopsies of 85 breast cancer patients over a period of 7 years of follow-up, and found no clear correlation between EBV and tumour characteristics (Marrão et al., 2014).

In the last four decades, EBV has been strongly linked to many autoimmune diseases. For example, the autoimmune disease systemic lupus erythematosus (SLE) has been linked with raised EBV-specific antibody titre and EBV viral load (Evans et al., 1971; Moon et al., 2004). However, a direct role of EBV in the development of SLE has not been established. Similarities between antibodies against EBNA1 and autoantibodies seen in SLE might indicate EBV as a causative agent (Ascherio and Munger, 2015).

Another example of an EBV-related autoimmune diseases is multiple sclerosis (MS), a chronic autoimmune disease that causes damage to the nerve cell sheaths, and affects nearly 6000 individuals every year across the UK (Mackenzie et al., 2014). The risk of developing MS was shown to increase substantially with the level of EBV antibody titre prior to the onset of the disease (Lunemann et al., 2007). Also, MS risk is 2.3 fold higher in patients with a history of infectious mononucleosis (Ascherio and Munger, 2015). The role of EBV in MS is unclear despite the fact that MS mostly occur in EBV- positive individuals with only a few exceptions where cases were reported in EBV-negative children. In addition, MS is associated with many risk factors such as the genetic alleles of HLA-DR15, vitamin D deficiency, and smoking (Ascherio and Munger, 2015).

An association between EBV and rheumatoid arthritis (RA) was proposed when sera from RA patients reacted specifically with EBV-positive B cell cultures (Alspaugh et al., 1978). RA is characterised by chronic inflammatory polyarthritis that causes damage to the synovial joints. The incidence rate of RA is around 1% worldwide and 2-4 times more prevalent in women compared to

men (Costenbader and Karlson, 2006). The sequence similarities between EBV-encoded proteins such as gp110 and the  $\beta$ -chain of human leukocyte antigen (HLA)-DR4 might play a critical role in RA development. In common with the other diseases, EBV plasma load is higher in RA patients than in healthy controls and EBV-specific cytotoxic CD8 T cells are commonly seen in affected joints tissue (Costenbader and Karlson, 2006). The development of erosive RA in humanised mice infected with EBV is also a strong indication of the role of EBV in RA (Kuwana et al., 2011).

In general, EBV is a plausible causative agent in autoimmune diseases. However, it is difficult to associate any single autoimmune disease with EBV. Like any other infectious agent, EBV might trigger an inflammatory reaction that leads to such diseases. EBV primary infection leads to the increased expression of several inflammatory cytokines such as IL-1 $\beta$ , TNF- $\alpha$ , and IL-6. B cell transformation by EBV initiates CD4<sup>+</sup>T cells immune response that results in the expansion of EBV-specific CD8<sup>+</sup> cells (Costenbader and Karlson, 2006). These diseases may therefore be the result of interactions between genetic predisposition factors and environmental factors that lead to an inflammatory reaction against self-antigens.

#### 1.4 Zta (BZLF1)

The EBV immediate-early transcription factor Zta (also called BZLF1, Zta, ZEBRA, Z, and EB1) triggers the switch from latent to lytic infection (Countryman and Miller, 1985; Takada et al., 1986). The term Zta will be used throughout to refer to Zta protein.

Zta expression is encoded by three exons in the *BZLF1* gene and derived from the *BZLF1* promoter (Zp). *BRLF1* (Rta) gene is located upstream of Zp and transcribed from the *BRLF1* promoter (Rp) (**Figure 1-5 (A)**). A longer alternative splice variant of the *BRLF1* mRNA produces a fusion transcript of BZLF1/BRLF1. The fused protein, known as RAZ, interferes with Zta transcriptional activation (Longnecker et al., 2013).

As discussed in **Section 1.3.4.3**, Zta is essential for the reactivation of latency in EBV infected cells (Feederle et al., 2000). Moreover, Zta orchestrates the

regulation of lytic cycle by directly regulating target viral and cellular promoters, hijacking various cellular pathways and initiating viral DNA lytic replication.

#### 1.4.1 Zta structure and DNA binding

Zta is a basic leucine zipper (bZIP) protein which belongs to the AP-1 protein family. This family includes the human cellular proteins c-Jun, c-Fos and ATFs as well as the yeast GCN4 (Farrell et al., 1989; Lieberman and Berk, 1990). Zta shares homology in its basic domain with the AP-1 members, in particular c-Fos (Farrell et al., 1989). However, Zta contains 4 atypical heptad repeats (unlike other bZIP proteins) (Chang et al., 1990; Flemington and Speck, 1990c; Kouzarides et al., 1991). Zta is 245 amino acids (aa) in length and has a molecular weight of 35 KDa (Marschall et al., 1989; Seibl et al., 1986). This protein is encoded by three exons; each exon encodes a separate functional domain (Flemington et al., 1992; Lin and Flemington, 2010). Zta forms a homodimer and cannot dimerise with other cellular bZIP proteins to form heterodimers (Chang et al., 1990). Amino acid residues between 168-202 show similarity with c-Fos bZIP domain (Taylor et al., 1991).

As shown in **Figure 1-4**, Zta can be divided into three major domains: amino-terminal transactivation (TA) domain (residues 1-167), basic DNA binding domain (DBD) (residues 175-197) and leucine zipper dimerization domain (DIM) (residues 197-221) (Flemington et al., 1992; Lin and Flemington, 2010).

The three-dimensional structure of the Zta fragment (residues 175–237) harbouring two single point mutations (S186A and C189S) and bound to DNA (AP-1 site) was solved at 2.25 Å resolution (**Figure 1-4 (A)**) (Petosa et al., 2006). This has demonstrated the unexpected structure of Zta where the C-terminal residues fold back to stabilise the coiled-coiled dimerisation domain (Petosa et al., 2006; Schelcher et al., 2007; Sinclair, 2006).

Zta has been subjected to large-scale mutagenesis studies which have revealed some of the structural features as well as the critical amino acid residues for Zta function. The TA domain is essential for Zta transcriptional activation, specifically, the region between residues 27-78 (Deng et al., 2001;

Flemington et al., 1992). The TA domain is also important for viral DNA replication; specifically, to interact and recruit the viral helicase (BBLF4). Zta with four point mutations (residues 22,26,74 and 75) failed to induce viral lytic DNA replication (Liao et al., 2001). These same Zta residues are required for Zta interaction with TFIIA-TFIID and the CREB-binding protein (CBP), and subsequently the stimulation of CBP histone acetyltransferase activity (Deng et al., 2001; Lieberman et al., 1997).

The smallest region required for specific DNA binding is between residues 172-227. However, the region between residues 141-245 is needed to give the same binding affinity seen in full-length Zta (Taylor et al., 1991). The atomic structure of Zta indicated that amino acid residues 182,185, 186 and 190 were in direct contact with the AP-1 site (Petosa et al., 2006). Zta with mutations in residues 178 to 180 or 187 to 189 failed to bind DNA (Taylor et al., 1991). In addition, the following Zta mutants Y180A, Y180E, or K188A were reported to affect DNA lytic replication (Heston et al., 2006). Deletion of the region between residues 153-167 was shown to affect Zta binding affinity *in vitro* (Flemington et al., 1992). Several proteins (e.g. NF- $\kappa$ B, P53, Ku80 and EBV BMRF1) interact with the C-terminal region (Chen et al., 2011; Dreyfus et al., 2011; Zhang et al., 1996). The C-terminal tail folds back against the zipper region which makes it important for stable dimerisation (McDonald et al., 2009; Petosa et al., 2006; Sinclair, 2006). Mutational analysis also showed this region is critical for binding with 53BP1 as well as DNA replication (Bailey et al., 2009a; McDonald et al., 2009).

Zta is also highly phosphorylated, and many crucial phosphorylation sites have been identified. The serine at 186 is phosphorylated by protein kinase C. The serine residue at 186 is found in Zta but not in the AP-1 protein c-Fos and it is vital for Zta binding at methylated ZREs in the BRLF1 promoter (Rp) but not binding to consensus AP-1 site (Adamson and Kenney, 1998; Bhende et al., 2005; El-Guindy and Miller, 2004). The Zta S186A mutant failed to reactivate lytic cycle (Francis et al., 1997). Similarly, the Zta C189S mutant failed to bind methylated ZREs in Rp (Karlsson et al., 2008a). Also, S167, S173 are phosphorylated by Casein Kinase 2 (CK2). This phosphorylation is important for



the ability of Zta to repress Rta activation of the viral *BLRF2* gene (El-Guindy and Miller, 2004). The Serine at 209 is phosphorylated by the viral kinase BGLF4; this phosphorylation mediates the interaction between Zta and BGLF4. This interaction results in the repression of Zta autoregulation at Zp (Asai et al., 2006; Asai et al., 2009).

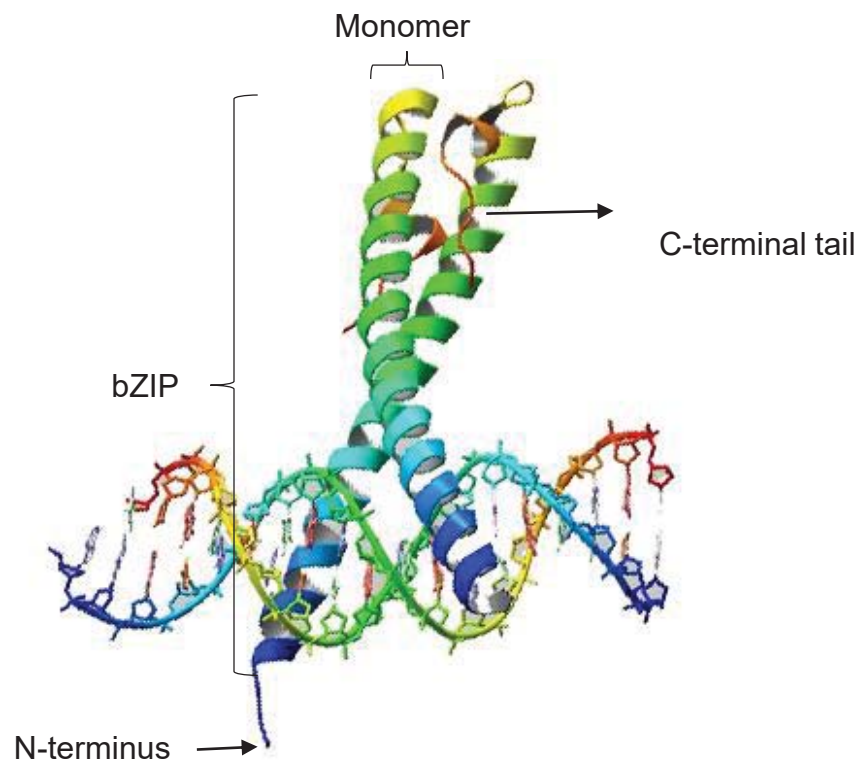
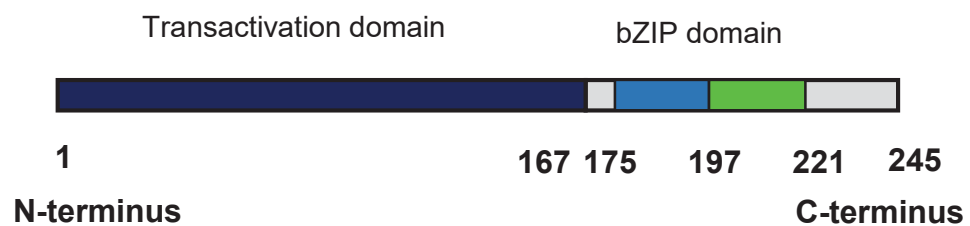
A mutation introduced at C171 affected Zta DNA binding ability and impaired its capacity to initiate lytic cycle replication and transcription activation at several promoters such as BMRF1 and BRLF1 (Rp) (Wang et al., 2006). Similarly, a mutation at C189 impaired the ability of Zta to induce or to mediate cell cycle arrest and is required for Zta transcription activation function (Schelcher et al., 2005).

Zta is also known to be sumoylated at K12 which inhibits its transcriptional activity. A mutation in this site was shown to affect EBV DNA replication but not Zta transcriptional function (Adamson and Kenney, 2001; Deng et al., 2001; Hagemeier et al., 2010).

Zta binds to the AP-1 canonical site (TGAGTCA) as well as many other sites related to AP-1 site, which are known as Zta response elements (ZREs) (Hardwick et al., 1988; Lieberman et al., 1990). Zta preferentially binds to methylated ZREs which is critical for activating the heavily methylated (silenced) EBV genome during latency (Bhende et al., 2004).

ZREs are classified into unmethylated and methylated ZREs (Bergbauer et al., 2010; Karlsson et al., 2008b). The unmethylated ZREs lack a CpG motif (class I ZREs). The methylated ZREs contain a CpG motif and are further classified into ZREs where Zta binding is enhanced by methylation (class II ZRE) or ZREs where Zta only recognises the methylated site (class III ZRE) (Karlsson et al., 2008b).



**A****B**

**Figure 1-4 Zta (BZLF1) structure.** Zta is a basic leucine zipper (bZIP) protein which belongs to the AP-1 protein family. **(A)** The three dimensional structure of the Zta fragment (residues 175–237) harbouring two single point mutations (S186A and C189S) bound to DNA (AP-1 site) (Petosa et al., 2006). **(B)** A schematic illustration of Zta protein showing the transactivation domain (dark blue), the bZIP domain which consists of the basic DNA binding domain (blue) and the dimerisation domain (green). The unstructured C-terminal tail is shown in light blue.

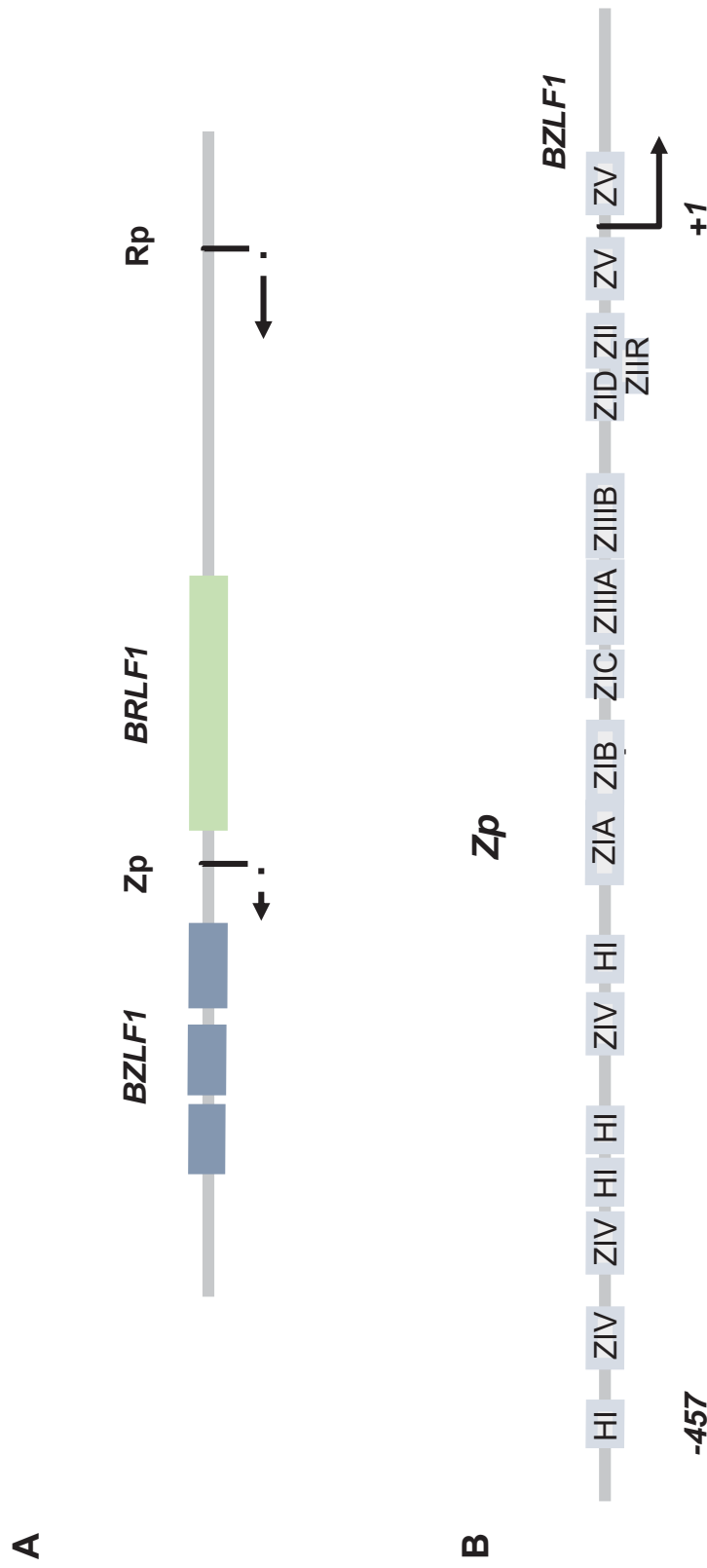
### 1.4.2 Zta promoter (Zp)

Zta regulation at the transcriptional level is mediated by several cellular factors that are involved in either negatively or positively regulating its promoter. The Zta promoter (Zp) is divided into several cis - regulatory regions numbered ZI to ZV (**Figure 1-5 (B)**) (reviewed in (El-Guindy et al., 2013; Kenney and Mertz, 2014)).

ZI has four regulatory elements (ZIA-ZID) that are regulated by SP1/KLF and MEF2 transcription factors. MEF2D acts as a Zp repressor during latency in B cells and binds ZIA, ZIB, and ZID elements. MEF2D repression is mediated by attracting histone deacetylation complexes (HDACs) to the promoter (Gruffat et al., 2002). Upon BCR receptor activation, MEF2D, which also bind ZII, becomes dephosphorylated by the cyclosporin A-sensitive calcium-mediated signalling pathway. Dephosphorylated MEF2D associates with histone acetylases, therefore, activating Zp (Bryant and Farrell, 2002). The ZII domain contains binding site for AP-1 and CREB proteins, which are important for Zp activation through AP-1 proteins. Both ZI and ZII are important for Zp activation through BCR crosslinking as well as TPA induction (Kenney and Mertz, 2014).

ZIII has two ZRE elements for Zta binding (ZIIIA and ZIIIB) that mediate Zta autoregulation of Zp (Flemington and Speck, 1990a). Zta activates the promoter through binding to ZII and ZIIIB by associating with C/EBP $\alpha$  (Wu et al., 2004). SMAD proteins also play a role in Zp activation through five SMAD-binding elements (SBEs), located throughout Zp (El-Guindy et al., 2013).

ZIV sites are involved in Zp silencing and bound by Zinc finger protein YY1 (Kenney and Mertz, 2014). The ZV and ZV' elements, located around the Zp transcription start site, are bound by E-box binding proteins ZEB1 and ZEB2 (also called SIP1). These two proteins synergistically mediate Zp silencing (Ellis et al., 2010; Kraus et al., 2003). Other elements (HI and ZIIR elements) are also associated with Zp repression. The HI elements are bound by E-box binding protein E2-2, while the ZIIR element is bound by a protein that remains to be identified (Kenney and Mertz, 2014).



**Figure 1-5 BZLF1 locus and promoter.** (A) Zta is encoded by three separate exons that forms the *BZLF1* gene (blue boxes) and transcribed from the *BZLF1* promoter (Zp). The immediate-early lytic gene *BRLF1* (Rta) is located upstream of *BZLF1* and transcribed from the *BRLF1* promoter (Rp). (B) A schematic illustration of Zp showing the various regulatory elements within the promoter. The promoter (Zp) is divided into several cis - regulatory regions numbered ZI to ZV. Image adapted from (El-Guindy et al., 2013; Kenney and Mertz, 2014).

### 1.4.3 Zta protein interactions and functional diversity

Zta is widely involved in many cellular pathways either directly or by regulating other viral and cellular genes. Also, it interacts and co-associates with many viral and cellular proteins (reviewed in (Sinclair, 2003)). For example, Zta interacts with the tumour suppressor protein p53 and interferes with its transcriptional function, preventing cell apoptosis (Mauser et al., 2002; Ning et al., 2003; Zhang et al., 1994). The anti-apoptotic role of Zta is further mediated by activating *BHRF1* which encodes a viral BCL2 homolog (Ragoczy and Miller, 1999) as well as activating human IL-10 (Mahot et al., 2003b).

Zta was also shown to induce IL-13 by direct binding to the IL-13 promoter. The induction of IL-13 is believed to be important in the early stage of EBV infection to promote proliferation and growth of EBV-immortalised lymphoblastoid cell lines (LCLs) (Tsai et al., 2009). Also, Zta is known to upregulate IL-8 directly through binding to two ZREs within IL-8 promoter (Hsu et al., 2008) and indirectly through EGR-1 activation (Heather et al., 2009; Hsu et al., 2008). IL-8 mediates infiltration of lymphocytes which indicates a possible role for Zta in EBV-positive lymphoma development and maintenance (Hsu et al., 2008). Zta is also involved in the regulation of several cytokines including inducing cellular cytokines that are involved in cell growth and proliferation such as TGF- $\beta$ , and VEGF (reviewed recently in (Murata and Tsurumi, 2014)).

The cellular CREB-binding protein (CBP) protein was shown to interact with Zta (Adamson and Kenney, 1999; Chen et al., 2001; Deng, 2003; Zerby et al., 1999). This interaction requires the Zta dimerization domain (Adamson and Kenney, 1999). Furthermore, Zta is known to activate the cellular stress mitogen-activated protein (MAP) kinases (p38 and JNK) which result in ATF2 phosphorylation and activation. ATF2 plays an important role in activating Zp through ZII region of the promoter (Adamson et al., 2000).

Zta interacts and inhibits NF $\kappa$ B subfamily protein p65 as well as interferon regulatory factor 7, which helps in viral immune evasion. The role of Zta in immune evasion and transcriptional regulation is discussed further in the next sections.

The mitochondrial single-stranded DNA-binding protein (mtSSB) was found to co-immunoprecipitate with endogenous Zta during lytic cycle reactivation in the EBV-positive BL Mutu I and in LCLs. mtSSB knockdown affected Zta-induced viral replication but not its transcription regulation ability. A single point mutation in the Zta DNA binding domain (C189S) prevented mtSSB association with Zta (Wiedmer et al., 2008).

Additionally, 53BP1, a DNA damage repair protein, associates with Zta through a region containing two BRCT domains (Bailey et al., 2009a). The C-terminus of both proteins was shown to be important for this interaction. Reduced viral replication was observed when 53BP1 was knocked down indicating its importance for viral replication (Bailey et al., 2009a).

Zta interaction with Oct-2 inhibits Zta binding to lytic gene promoters and hence prevents lytic genes reactivation. The interaction between Zta and Oct-2 is mediated by the POU domain of Oct-2 and the DNA-binding/dimerization domain of Zta (Robinson et al., 2012).

Additionally, high expression of Zta is known to disturb promyelocytic leukaemia bodies (PML). These nuclear structures contain many proteins and thought to be involved in many cellular functions such as apoptosis, MHC I presentation and DNA damage repair (Adamson and Kenney, 2001; Bell et al., 2000; Kenney, 2007).

Zta hijacks the cell cycle to promote favourable conditions for viral lytic cycle DNA replication. Zta appears to mediate G0/G1 growth arrest and blocks G1/S transition. Zta facilitates this arrest by inducing cyclin-dependent kinase inhibitors, p21 and p27 as well as tumor suppressor protein p53. Furthermore, Zta reduces the expression of cyclin A and c-Myc (Cayrol and Flemington, 1996a; Cayrol and Flemington, 1996b; Kudoh et al., 2003; Rodriguez et al., 2001; Wu et al., 2003; Zhou et al., 2005). Zta induction of p21 expression involves the direct interaction with CCAAT/Enhancer Binding Protein C/EBP- $\alpha$  protein. Zta alone did not activate p21 promoter, but activation was reported in

the presence of C/EBP- $\alpha$  through a C/EBP- $\alpha$  binding sites in p21 promoter (Wu et al., 2003).

Zta is essential for viral DNA replication during lytic cycle. It binds several ZREs in OriLyt and recruits viral replication proteins to OriLyt as previously described in **Section 1.3.4.3**. Zta interacts with most viral replication proteins except for BALF2 (El-Guindy et al., 2007; McKenzie and El-Guindy, 2015).

#### 1.4.4 Zta transcriptional regulation function

Zta induces transcription of its target promoters mainly by direct binding to Zta response elements (ZREs) including AP-1 sites within regulatory elements (Lieberman and Berk, 1990; Lieberman et al., 1990) as described previously in **Section 1.4.3**. Additionally, Zta, through its transactivation domain, recruits and stabilises the general transcription factors TFIID and TFIIA at the core promoter to facilitate transcription (Chi and Carey, 1993; Lieberman, 1994; Lieberman and Berk, 1991, 1994; Lieberman et al., 1989; Lieberman et al., 1997). Zta is known to interact with coactivators CREB-binding protein (CBP) and its related protein P300 to stimulate histone acetyltransferase activity, and facilitate chromatin acetylation (Adamson and Kenney, 1999; Chen et al., 2001; Deng, 2003; Zerby et al., 1999). Zta is thought to be involved in either the direct or indirect regulation of several viral and cellular genes. For example, Zta interacts synergistically with the cellular oncogene c-myc to activate *BMRF1* promoter which requires Zta binding at the promoter (Kenney et al., 1992). A primary target for Zta is the viral immediate-early gene *BRLF1* (Sinclair et al., 1991). Zta binds preferentially to methylated sites on *BRLF1*, a unique feature that is essential for the induction of lytic cycle by Zta (Bhende et al., 2004, 2005).

Zta is known to directly activate several lytic genes such *BMRF1* (Holley-Guthrie et al., 1990), *BALF2* (Hung and Liu, 1999), *BHLF1* (Lieberman et al., 1989) and *BNLF2a*, as will be shown in **Chapter 5**. Furthermore, the viral lytic genes (*BSLF1*, *BSLF2*, *BALF2*, *BALF5*, *BBLF2/3*, and *BBLF4*) are shown to be direct targets to Zta (Bergbauer et al., 2010).

Zta is known to autoregulate its own promoter (Zp) through direct binding to two adjacent ZREs within ZIIIA and ZIIIB regions of the promoter (Flemington and Speck, 1990a). The immediate-early protein Rta (BRLF1) is also activated by Zta (Bhende et al., 2004; Sinclair et al., 1991).

Zta synergises with Rta through cellular MCAF1 and RanBPM to activate several viral promoters such as *BHLF1*, *BHRF1*, and *BRLF1* (Chang et al., 2010; Liu and Speck, 2003; Ragoczy and Miller, 1999; Yang et al., 2015).

Zta, through its c-terminal region interacts with Ku80, a member of the DNA-dependent protein kinase (DNA-PK) complex, to enhance the activation of *BHLF1* promoter. This activation is mediated through Zta binding to ZREs within the promoter (Chen et al., 2011). The Ku protein is a heterodimer consisting of Ku70 and Ku80 and it is a known binding partner of Zta (Bailey et al., 2009a; Wiedmer et al., 2008).

A number of cellular genes have also been identified as direct targets for Zta regulation (reviewed in (Chen et al., 2009; Sinclair, 2003). Activation of the host gene *EGR-1* is mediated by Zta indirectly and directly through binding to ZREs within the *EGR-1* promoter (Chang et al., 2006; Heather et al., 2009). This activation was enhanced at least 10-fold when the Zta binding site was methylated (Heather et al., 2009).

A different example of Zta regulation is seen in the host *CIITA* gene. This important gene is constitutively expressed and encodes a transcription factor which is important in inducing MHC II expression (Germain, 1994). Unlike the previous promoters, *CIITA* is known to be downregulated by Zta; this is mainly mediated through a Zta binding site identified in the PIII region of the *CIITA* promoter. Although two ZREs ~1000 bp apart were identified by sequence analysis, Zta binding and regulation was dependent on one site only (Li et al., 2009). Moreover, the TA domain of Zta was shown to reduce the activity of this promoter indicating a possible DNA binding-independent mechanism involved in *CIITA* downregulation by Zta. A shorter region of the promoter lacking Zta

binding sites also showed reduced activity upon Zta transfection (Balan et al., 2016).

Several other cellular proteins are either positively or negatively regulated by Zta at the transcriptional level including human IL-10, c-Fos, and c-Jun (Flemington and Speck, 1990b; Mahot et al., 2003a; Sato et al., 1992).

### **1.5 EBV immune evasion**

The co-existence of EBV with its host requires the virus to stay invisible to the sophisticated host immune system. Immediately upon viral infection, the host innate immune response is triggered. This response is facilitated by the pattern recognition receptors (PRRs) that are expressed in several immune and non-immune cells. Their function is to survey for any pathogen ligands such as nucleic acids, carbohydrates and proteins (commonly referred to as pathogen-associated molecular patterns (PAMPs)). The PRRs are either cell surface receptors or unbound intracellular receptors (cytoplasmic). There are five families of PRRs including the Toll-like receptors (TLR) and RIG-I-like receptors. Activation of the PRRs expressed in the immune cells such as dendritic and macrophages cells mediates the secretion of many antiviral cytokines including interferons type I (IFN I). The biology and function of each receptor are reviewed in (Brubaker et al., 2015).

The adaptive immune system also plays a role in blocking any viral infection. CD8 T cells are thought to be the major player in controlling primary viral infections. These specialized cells recognise viral peptides presented by MHC I molecules on the surface of infected cells. CD8 T cells have a cytotoxic function by secreting of cytolytic proteins such as Granzyme B and perforin, as well as cytokines such as IFN $\gamma$  and TNF $\alpha$ . CD4 T subtype cells help in secreting various inflammatory cytokines and promote B cell survival and proliferation (Reviewed in (Zuniga et al., 2015)).

Viruses including EBV have evolved to have various immune evasion mechanisms to escape the host innate and adaptive immune responses. These



mechanisms are usually dependent on viral life cycle stage. For example, during the entry of the virus, the EBV envelope glycoprotein gp42, which facilitate the viral entry and fusion to the host cell, binds to MHC II. This is thought to block MHC II interaction with CD4 T cells (Ressing et al., 2003). Moreover, the blockage of MHC II is mediated by a cleaved soluble form of gp42 (Ressing et al., 2005b).

During pre-latency as well as lytic cycle, the expression of EBV lytic genes such as *BZLF1*, *BCRF1* and *BNLF2a* are critical for immune evasion. Zta (*BZLF1*) interferes with the production of INF $\alpha$ 4 and INF $\beta$  by negatively regulating the interferon regulatory factor 7 (IRF7) (Hahn et al., 2005). Zta interferes with p65 transcription activation, a subunit of NF- $\kappa$ B complex. This inhibits NF- $\kappa$ B signalling pathway and affects the regulation of many immune response genes triggered by this pathway (Morrison and Kenney, 2004). Furthermore, Zta directly downregulates the transcription of MHC class II trans activator (CIITA), a gene that promotes expression of MHC-II molecules (Balan et al., 2016; Li et al., 2009). In addition, Zta downregulates surface CD74, a chaperone for MHC II antigen presentation (Zuo et al., 2011b) On the other hand, BNLF2a inhibits viral peptide presentation on MHC I by inhibiting peptide loading into the transporter associated with antigen presentation (TAP) as will be discussed in the next section.

BGLF5 and BILF1 are two other lytic cycle proteins which are involved in interfering with MHC I (Ressing et al., 2008; Zuo et al., 2009). BGLF5 inhibits cellular protein synthesis during lytic cycle. In EBV, this global protein synthesis shutoff is mediated by the viral BGLF5 which has not only DNase activity (Lin et al., 1995; Stolzenberg and Ooka, 1990) but RNase activity that affects global cellular mRNA stability (van Gent et al., 2011). This is thought to affect the protein synthesis of many cellular proteins involved in immune response such as TLR9 (van Gent et al., 2011) as well as the synthesis of HLA class I and II molecules (Ressing et al., 2008; Rowe et al., 2007). In a more recent report, the role of BGLF5 in reducing MHC I was shown to be less important as the downregulation of BGLF5 had minimal effects on MHC I presentation (Quinn et al., 2014).

BILF1 encodes a constitutively active G-protein-coupled receptor (GPCR) (Paulsen et al., 2005). The downregulation of MHC I by BILF1 is not associated with its GPCR function but it is due to the interaction and interference with HLA I transport to the cell surface as well as increased degradation of these molecules (Zuo et al., 2011a; Zuo et al., 2008). Furthermore, BDLF3, a viral glycoprotein (Nolan and Morgan, 1995) promotes the ubiquitination of both MHC molecules to facilitate their proteasomal degradation (Quinn et al., 2016). It is worth noting that MHC I targeting by EBV proteins is believed to be phase dependent. For example, the role of BNLF2a in reducing peptide presentation on MHC I affects the presentation of immediate-early and early lytic genes. In contrast, BILF1 affects antigen recognition of early and late lytic genes (Quinn et al., 2014).

Furthermore, BCRF1, the host IL-10 homologue affects T cell activation by inhibiting various immune response modulators such as MHC I and MHC II complexes, as well as the ICAM 1 receptor (a receptor that facilitates the interaction between leukocytes and epithelial cells) (Salek-Ardakani et al., 2002).

During latency, EBNA1, which is expressed at all latency phases, is known to have an internal Gly-Ala repetitive sequence that leads to the prevention of its own peptide presentation on MHC I (Dantuma et al., 2002; Levitskaya et al., 1995; Levitskaya et al., 1997). Various EBV latent proteins and miRNAs have been reported to be involved in immune evasion mechanisms; mostly by downregulating pro-inflammatory proteins and pathways (reviewed in (Ressing et al., 2015)).

### **1.5.1 The role of BNLF2a in immune evasion**

Major histocompatibility complex class I (MHC I) molecules play a central role in host immune recognition of foreign antigens. The primary role of MHC I is to present short peptides processed within the cell to CD8 T cells. Circulating cytotoxic T lymphocytes (CTL) (CD8 T cells) differentiate between self and non-self peptides. In a virus-infected cell, processed peptides from virus proteins are

presented on MHC I. Such cells are flagged by CTL as foreign antigens and are destroyed by CTLs. Thus, CTL immune response is dependent on efficient viral peptide presentation by MHC I in infected cells. The antigen processing starts with the degradation of intracellular proteins into very short peptides (approximately 8-16 aa). Protein degradation is mediated by the proteasome in the cytosol and resulting peptides are then transported into the endoplasmic reticulum (ER) where they associate with the transporter associated with antigen processing complex (TAP) at the ER membrane. Subsequently, peptides bound to TAP are translocated to the MHC I molecules. Mature MHC I molecules loaded with peptides exit the ER to be presented on the cell surface. TAP is a heterodimeric protein consisting of TAP1 and TAP2 subunits, both containing an N-terminal transmembrane domain and a C-terminal cytosolic nucleotide binding domain. The latter domain is associated with the peptide binding. Moreover, ATP molecules binding to TAP is essential for peptide transportation (reviewed in (Neefjes et al., 2011)).

Several herpesviruses have developed various mechanisms to target the antigen presentation process to evade any immune response. For example, the HSV ICP47 protein is known to target TAP molecules by direct binding. CMV US6 is also known to inhibit the ATP binding by targeting TAP molecules (Røder et al., 2008).

Similarly, EBV BNLF2a, a 60 amino acids protein (6.5 KDa) expressed from the *BNLF2a* (ED-L2) promoter (Hudson et al., 1985), was shown to interfere with both peptide and ATP binding to TAP molecules within the ERs compartments (Hislop et al., 2007). The BNLF2a protein is a tail-anchored protein composed of a hydrophilic N-terminal region and a C-terminal hydrophobic part (20 amino acids). The C-terminal tail anchor mediates ER membrane association and retention, while its cytosolic N-terminus inhibits peptide and ATP binding to TAP (Wycisk et al., 2011).

B cells infected with a mutant EBV, harboring a deletion in the *BNLF2a* gene, showed reduced recognition of viral immediate-early and early antigens by specific CD8 T cells. However, the recognition of late lytic antigens was not

affected by *BNLF2a* deletion when compared to wildtype EBV (Croft et al., 2009). It is worth noting that BNL2a expression was detected in nearly 50% of gastric carcinoma biopsies analysed by (Strong et al., 2015). Low expression of Zta was also detected in these samples. However, there was no significant correlation between the expressions of both proteins suggesting that BNL2a expression is independent of Zta (Strong et al., 2015).

Although, the role of BNL2a results in a reduced viral peptide presentation on the cell surface, recent findings showed that a peptide from BNL2a C-terminal region is presented on MHC I and recognized by CD8 T cells (Bell et al., 2009). This further supports that a collaborative role for EBV proteins in immune evasion is critical for viral immune evasion as described in the previous section.

## 1.6 Project Aims

A genome-wide ChIP-Seq experiment to detect Zta binding was carried out in BL Akata cells previously in our lab. This has identified binding sites for Zta across the viral and cellular genome. In this project, we aimed to investigate conservation of Zta binding across the viral genome in different cell lines including non-B cells.

The genome-wide analysis of Zta binding across the host genome also showed that Zta was associated with distal enhancer elements. Therefore, we asked the question of whether Zta is able to enhance activation of a minimal promoter through distal ZRE element using luciferase reporter constructs.

Zta binding was also observed in the ChIP-Seq data at the viral early lytic gene *BNLF2a* suggesting a role for Zta in regulating this promoter. We set out to characterise Zta regulation of this important gene for viral immune evasion and use it as a model to broaden our understanding of Zta transcriptional regulation of the viral early lytic genes.

## Chapter 2 Materials and Methods

### 2.1 Materials, and reagents

#### 2.1.1 DNA constructs (plasmids)

Plasmid Name	Generated by/ Reference	Backbone plasmid	Application
<b>pcDNA3</b>	Invitrogen	pcDNA3	Control for Zta expression
<b>His-Zta (Full-length Zta)</b>	Previously in the lab (Bailey et al., 2009a)	pcDNA3	Zta expression
<b>His-GFP- bZIP Zta (Zta 168-245 aa)</b>	Myself with the help of Dr Louise Bird at Oxford Protein Production Facility (OPPF)	pOPINN-GFP (OPPF)	Protein purification and DNA binding
<b>PGL3-Control luciferase</b>	Promega	PGL3	Zta enhancer element analysis
<b>MinC-Null (CIITA -214+53)</b>	Previously in the lab by Nicolae Balan (Balan et al., 2016)	PGL3-Control	
<b>MinC-ZREs</b>	Undergraduate student: Renu Gurung	PGL3-Control	
<b>MinP-Null</b>	Myself	PGL3-Control	
<b>MinP-ZREs</b>	Myself	PGL3-Control	
<b>pCpGL luciferase</b>	(Klug and Rehli, 2006)	pCpGL	BNLF2a promoter analysis
<b>BNLF2a 1-5</b>	Previously in the lab by Kay Osborn	pCpGL	
<b>BNLF2a Null</b>	Previously in the lab by Kay Osborn	pCpGL	
<b>BNLF2a 4-5</b>	Myself	pCpGL	
<b>BNLF2a 1-3</b>	Myself	pCpGL	
<b>BNLF2a 2-5</b>	Myself	pCpGL	
<b>BNLF2a 1&amp;3-5</b>	Myself	pCpGL	
<b>BNLF2a 1-2&amp;4-5</b>	Myself	pCpGL	
<b>BNLF2a 1</b>	Myself	pCpGL	
<b>BNLF2a 2</b>	Myself	pCpGL	
<b>BNLF2a 3</b>	Myself	pCpGL	
<b>BNLF2a 1&amp;2</b>	Myself	pCpGL	
<b>BNLF2a 1&amp;3</b>	Myself	pCpGL	
<b>BNLF2a 2&amp;3</b>	Myself	pCpGL	
<b>BNLF2aTATA M</b>	Myself	pCpGL	
<b>BNLF2aKLF4 M</b>	Myself	pCpGL	

<b>BNLF2a 1FL M</b>	Myself	pCpGL	
<b>BNLF2a 2FL M</b>	Myself	pCpGL	
<b>BNLF2a 3FL M</b>	Myself	pCpGL	
<b>BNLF2a 4FL M</b>	Myself	pCpGL	
<b>BNLF2a 5FL M</b>	Myself	pCpGL	
<b>BNLF2a 1-5FL M</b>	Myself	pCpGL	

**Table 2.1: Plasmid DNA constructs used in various experiments.**

### 2.1.2 Cell lines

<b>Cell line name</b>	<b>Description</b>	<b>Reference</b>	<b>Application</b>
<b>LCL#3</b>	Lymphoblastoid B cells transformed using B95-8 EBV strain	(Sinclair et al., 1994b)	Zta ChIP-qPCR
<b>293-BZLF1-KO (also called 293 Z-KO)</b>	293 cells stably infected with recombinant EBV lacking BZLF1 gene (Semi-adherent cells)	(Feederle et al., 2000)	Zta ChIP-qPCR
<b>DG75</b>	EBV-negative Burkitt's lymphoma cells	(Ben-Bassat et al., 1977)	Zta enhancer element analysis, and BNLF2a promoter analysis
<b>HeLa cells</b>	Adherent epithelial cells	(Gey et al., 1952)	BNLF2a promoter analysis

**Table 2.2: Mammalian cell lines used in various experiments.**

### 2.1.3 Antibodies

<b>Antibody</b>	<b>Description</b>	<b>Application</b>	<b>Dilution</b>	<b>Supplier</b>
<b>Zta antibody (BZ1)</b>	Zta primary mouse monoclonal	Western Blot (WB)	1:500	Gift from Martin Rowe Lab, Birmingham Described in (Young et al., 1991c)
<b>Zta antibody (SCZ)</b>	Zta primary goat, polyclonal	ChIP/EMSA	10 µg/IP 2 µg /EMSA reaction	Santa Cruz Biotech (sc-17503)
<b>Normal goat IgG</b>	Goat	ChIP (control antibody)	10 µg/IP	Santa Cruz Biotech (sc-2028)

<b>ZEB1</b>	Goat polyclonal	EMSA/WB	2 µg/EMSA reaction	Santa Cruz Biotech (sc-10572)
<b>ZEB2 (SIP1)</b>	Rabbit polyclonal	EMSA/WB	2 µg/EMSA reaction	Bethyl Laboratories A302-474A
<b>β-Actin</b>	Rabbit polyclonal	WB	1:5000	Sigma (A2066)
<b>IRDye 680CW</b>	anti-mouse secondary fluorescent Ab	WB	1:5000/1:10000	LI-COR
<b>IRDye 800CW</b>	anti-rabbit secondary fluorescent Ab	WB	1:5000/1:10000	LI-COR

**Table 2.3: Antibodies used in various experiments.**

## 2.1.4 Primers (oligonucleotides)

### 2.1.4.1 ChIP-qPCR primers

EBV gene target	qPCR primers pair for ChIP (5' to 3')	PCR product genome coordinates (NC_007605.1)
<b>BNRF1</b>	F <sup>(1)</sup> TGTGACACCAACAGGTGTTGCCTTG	1501-1570
	R <sup>(1)</sup> ACCCCAAAGAGGGCAAAGCCTAC	
<b>BCRF1</b>	F <sup>(1)</sup> GGGAGGTACATGTCCCCAGCATT	9253-9311
	R <sup>(1)</sup> CTGTGGACTGCAACACAACATTGCC	
<b>BFLF2</b>	F <sup>(2)</sup> ATCTGCAGCCAGGCCCTTAGCC	44985-45015
	R <sup>(2)</sup> CAAAACACGCTGCTGGATGTGCC	
<b>BLLF3</b>	F <sup>(2)</sup> TCCATCTGGGCACTTCTGACGCT	76463-76542
	R <sup>(2)</sup> CCGTCAGCAGCGTGTTTCAAA	
<b>BLLF2</b>	F <sup>(2)</sup> CTCAGTGACATGGAAGAGGTTG	77863- 77970
	R <sup>(2)</sup> CCAACCACACCTTAGGAGGA	
<b>BZLF1</b>	F <sup>(2)</sup> AGCCAAGGCACCAGCCTCCT	90943-91018
	R <sup>(2)</sup> TGCATGAGCCACAGGCATTGCT	
<b>BRRF1</b>	F <sup>(1)</sup> CCTGTTGTTTCGGAGAATGG	92638-92770
	R <sup>(1)</sup> AATTACAGCCGGGAGTGTG	
<b>BRLF1</b>	F <sup>(1)</sup> GGCTGACATGGATTACTGGTC	9389-93981
	R <sup>(1)</sup> TGATGCAGAGTCGCCTAATG	
<b>BKRF4</b>	F <sup>(3)</sup> CATTGCTCTCTGAGCGGTTA	98591-98667

	R <sup>(3)</sup> ACCAGATGCTTCTTGGAGTTG	
<b>BTRF1</b>	F <sup>(3)</sup> AGCTACGCAATCGGAGTCA	126981-127034
	R <sup>(3)</sup> GGAGGCTCAGTCTAGCAG	
<b>OriLyt-Flank</b>	F <sup>(1)</sup> GCGCAACAGTGCCACCAACC	50517-50590
	R <sup>(1)</sup> CAGGACCTGGCGGTAGTGCAG	
<b>OriLyt</b>	F <sup>(1)</sup> CAGCTGACCGATGCTCGCCA	41196/144115- 41323/144242
	R <sup>(1)</sup> ATGGTGAGGCAGGCAAGGCG	
<b>BNLF2A</b>	F <sup>(2)</sup> GGCCGTGGGGGTCGTCATCA	167728- 167759
	R <sup>(2)</sup> ACGCTGCTTTTGGGTTCTTCTGGT	
<b>82Kbp</b>	F CCTCCGAGCCGTGTGAAGCTC	82594-82682
	R CAAGAACCTCGGTGGCATCCTGG	

**Table 2.4: Primers used in ChIP-qPCR experiments. ChIP-qPCR for EBV target genes.**

(1) (Ramasubramanyan et al., 2012b). (2) (Ramasubramanyan et al., 2012a). (3) (Flower et al., 2011).

#### 2.1.4.2 EMSAs primers

DNA Probe Name (5' to 3')	Sequence	Fluorescent dye	Primer Purification
<b>BNLF2a ZRE1 F</b>	ACACCTGTCCT <u>TCGCTCA</u> TCTTTCCACA	5' IRDye 800	HPLC
<b>BNLF2a ZRE1 R</b>	TGTGGAAAGAT <u>TGAGCGA</u> GGACAGGTGT	None	Desalted
<b>BNLF2a ZRE2 F</b>	CACCTGTTGT <u>TGACACA</u> ATTCTTTGCGC	5' IRDye 800	HPLC
<b>BNLF2a ZRE2 R</b>	GCGCAAAGAAAT <u>TGTGTCA</u> ACAACAGGTG	None	Desalted
<b>BNLF2a ZRE3 F</b>	CTTCCATCTT <u>TGTGCCA</u> AATACACATTT	5' IRDye 800	HPLC
<b>BNLF2a ZRE3 R</b>	AAATGTGTATT <u>TGGCACA</u> AGATGGAAAG	None	Desalted
<b>BNLF2a ZRE4 F</b>	TCACCTTAAC <u>TGGCACA</u> CACTCCCTTA	5' IRDye 800	HPLC
<b>BNLF2a ZRE4 R</b>	TAAGGGAGTGT <u>TGTGCCA</u> GTTAAGGTGA	None	Desalted
<b>BNLF2a ZRE5 F</b>	TAAGCTACTAT <u>TGACTAA</u> CCTTTCTTTA	5' IRDye 800	HPLC
<b>BNLF2a ZRE5 R</b>	TAAAGAAAGGTT <u>AGTCA</u> TAGTAGCTTA	None	Desalted
<b>BNLF2aZRE2M F</b>	CACCTGTTGT <u>CCCCTTT</u> TTCTTTGCGC	5' IRDye 800	HPLC
<b>BNLF2aZRE2M R</b>	GCGCAAAGAA <u>AAAGGGG</u> ACAACAGGTG	None	Desalted



ZRE2 E-BOX Flank motif			
<b>BNLF2a</b> ZRE2, E-box F	AGGAAC <b>CACCT</b> GTTGTT <u><b>TGACACA</b></u> ATTCTT	5' IRDye 800	HPLC
<b>BNLF2a</b> ZRE2, E-box R	AAGAA <u><b>TGTGTCA</b></u> ACAAC <b>AGGTG</b> TTTCCT	None	Desalted
<b>BNLF2a</b> ZRE2, E-box <b>M</b> F	AGGAA <b>GTTTG</b> GTTGTT <u><b>TGACACA</b></u> ATTCTT	5' IRDye 800	HPLC
<b>BNLF2a</b> ZRE2, E-box <b>M</b> R	AAGAA <u><b>TGTGTCA</b></u> ACAAC <b>CAAAC</b> TTTCCT	None	Desalted

**Table 2.5: Oligonucleotides used to generate EMSA probes.**

The ZRE motif in each oligonucleotide is underlined. The motifs coloured in red represent the mutant motif.

### 2.1.5 Solutions and buffers

Method	Name	Composition	Purpose
<b>General molecular biology methods</b>	Tris-Borate-EDTA (0.5/1XTBE)	89 mM Tris (pH 7.6), 89 mM boric acid, 2 mM EDTA	DNA agarose/EMSA retardation gel running buffer
<b>General protein methods</b>	WB transfer buffer	25 mM Tris-HCl (pH 8.3), 192 mM glycine, 20% (v/v) methanol	WB
	PBS -Tween (PBS-T)	138 mM NaCl, 2.7 mM KCL, 10 mM Na <sub>2</sub> HPO <sub>4</sub> , 1.8 mM KH <sub>2</sub> PO <sub>4</sub> (pH 7.4), 0.1%(v/v) Tween-20	WB
	Ponceau S stain	0.1% (w/v) Ponceau S in 5% (v/v) acetic acid	Nitrocellulose membrane staining
<b>Chromatin immunoprecipitation-qPCR</b>	100 mM PMSF	Phenylmethylsulfonyl fluoride (174 mg) in 10 ml of 100% ethanol	Added to all ChIP buffer at a final concentration of 1mM
	IP dilution buffer	0.01%(v/v) SDS, 1.1%(v/v) Triton X-100, 1.2 mM EDTA, 16.7 mM Tris (pH 8), 167 mM NaCl	ChIP equilibration buffer
	SDS lysis	1% (v/v) SDS, 10 mM EDTA, 50	Chromatin

	buffer	mM Tris (pH 8)	preparation
	Cell lysis buffer	85 mM KCl, 0.5% (v/v) NP-40, 5 mM PIPES (pH 8)	Cell lysis
	Low salt buffer	0.1% (v/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA, 20 mM Tris (pH 8), 150 mM NaCl	ChIP washes
	High salt buffer	0.1% (v/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA, 20 mM Tris pH 8, 500 mM NaCl	ChIP washes
	Lithium chloride (LiCl)	250 mM LiCl, 1%(v/v) NP-40, 1%(v/v) Na-deoxycholate, 1 mM EDTA, 10 mM Tris (pH 8)	ChIP washes
	ChIP TE buffer	10 mM Tris, 5 mM EDTA,	ChIP washes
	Elution Buffer	10 mM Tris, 5 mM EDTA, 1 mM PMSF, 1%(v/v) SDS	ChIP Elution
<b>Protein purification</b>	Lysis Buffer pH 7.5	50 mM Tris-Cl pH 7.5, 500 mM NaCl, 0.5 mM TCEP, ~350 units/ml Benzonase Nuclease, 0.2%(v/v) Tween-20	His-GFP-bZIP Zta recombinant protein purification
	Equilibration buffer	50 mM Tris-CL pH 7.5, 500 mM NaCl, TCEP 0.5 mM	
	Washing Buffer	50 mM Tris pH 7.5, 500 mM NaCl, 30 mM imidazole.	
	Elution buffer	50 mM Tris pH 7.5, 500 mM NaCl, 150 mM imidazole.	
	Size exclusion elution buffer	50 mM Tris pH 7.5, 300 mM NaCl, 0.5 mM TCEP	
<b>HeLa cells nuclear extract</b>	Cytoplasmic extract buffer (CE)	10 mM HEPES pH 7.5, 60 mM KCL, 1 mM EDTA, 0.075% (v/v) NP40, 1mM DTT, 1 mM PMSF	EMSA
	Nuclear extract buffer (NE)	20 mM Tris-CL pH 8, 420 mM NaCL, 1.5 mM MgCL2, 0.2 mM EDTA, 1mM PMSF, 25% (v/v) glycerol	EMSA
<b>EMSA</b>	Annealing buffer	20 mM Tris-HCL pH 8.0, 1mM EDTA pH 8.0, 50 mM NACL	EMSA
	DTT	1 M DTT stock (1.5 g in 10 ml nuclease-free water)	EMSA

Table 2.6: Solutions and buffers.

### 2.1.6 Kits and reagents

General molecular biology methods	Kit/ reagent	Purpose	Supplier
	Maxi/Midi preps Kit	Large-scale plasmid DNA extraction and purification	Qiagen
	Mini-preps	Small-scale plasmid DNA extraction and purification	Qiagen
	T4 DNA ligase	Sub-cloning	Invitrogen
	5X T4 ligase buffer	Used in ligation reaction	Invitrogen
	Penicillin	Bacterial cells selection	Sigma
	Zeocin	Bacterial cells selection	Invitrogen
	HindIII restriction enzyme	Restriction digest	Roche
	BamHI restriction enzyme	Restriction digest	Roche
	10X Buffer H	Restriction digest	Roche
	HindIII-HF restriction enzyme	Restriction digest	NEB
	BamHI-HF restriction enzyme	Restriction digest	NEB
	KpnI-HF	Restriction digest	NEB
	Sall-HF	Restriction digest	NEB
	DpnI	Digesting methylated DNA	NEB
	CutSmart buffer	Restriction digest	NEB
	rAPid Alkaline phosphatase	Digested Plasmid DNA phosphatase	Roche
	DNA Hyperladder Bioline	Agarose gel DNA marker	Bioline
	5X Bioline DNA loading blue dye	DNA agarose gel	Bioline
	Gel Red	DNA agarose gel	Biotium
	Top10 competent cells	Plasmid DNA transformation	Invitrogen
	PIR1 competent cells	pCpGL derived constructs DNA transformation	Invitrogen
	Rosetta PlysS competent cells	Protein expression	Novagen
	Phusion Flash Master Mix	PCR	Thermoscientific

	AMPure XP magnetic bead	PCR purification	Beckman Coulter
	Bnzenase Nuclease	Protein purification	Sigma
	Luria Broth	Bacterial cultures	LifeSci school service department
	Power Broth™	Bacterial cultures (protein expression)	Molecular Dimensions
<b>General protein methods</b>	Odyssey blocking buffer	Diluting fluorescent secondary antibodies for WB	LI-COR
	Dried milk powder	WB membrane blocking	co-operative
	NuPAGE Bis-Tris SDS-PAGE (10% and 12%)	Protein gel (SDS-PAGE) electrophoresis	Invitrogen
	NuPAGE MOPS buffer	Protein gel (SDS-PAGE) electrophoresis	Invitrogen
	SeeBlue Pre-stained protein marker	Protein SDS-PAGE gel electrophoresis and western blots	Invitrogen
	2X protein sample buffer Lamelli	Protein SDS-PAGE gel electrophoresis	Sigma
	SimplyBlue SafeStain	SDS-PAGE staining	Invitrogen
<b>Tissue culture methods</b>	DMEM medium	Adherent cells culture	GIBCO, Invitrogen
	RPMI 1640 medium	Suspension cells culture	GIBCO, Invitrogen
	Foetal calf serum (FCS)	Cell culture medium supplementary	GIBCO, Invitrogen
	Penicillin, Streptomycin, L-Glutamine (PSG)	Cell culture medium supplementary	GIBCO, Invitrogen
	DPBS	Cell washing	GIBCO, Invitrogen
	Trypsin-EDTA (0.25%), phenol red	Adherent cell dissociation	GIBCO, Invitrogen
	Dimethyl sulfoxide (DMSO)	Cell freezing	Sigma
	Hygromycin B	293-BZLF1-KO cells selection	Invitrogen
	Effectene Kit	Adherent cells transfection	Qiagen

<b>Chromatin immunoprecipitation-qPCR</b>	Protein G Sepharose beads	Chromatin immunoprecipitation	Sigma
	Protein A Sepharose beads	Chromatin immunoprecipitation	Sigma
	Salmon sperm DNA	Sepharose beads blocking	Invitrogen
	GoTaq Syber green master mix	Real time quantitative PCR	Promega
<b>Luciferase reporter assay</b>	5x Passive Lysis Buffer (PLB)	Cell lysis	Promega
	Luciferase reporter assay kit	Luciferase measurement	Promega
<b>EMSA</b>	10X orange loading dye	EMSA	LI-COR
	5X binding buffer Consist of (50 mM Tris-HCl (pH 7.5), 5 mM MgCl <sub>2</sub> , 2.5 mM EDTA, 2.5 mM DTT, 250 mM NaCl, 0.25µg/ul poly (dI-dC), 20%(v/v) glycerol	EMSA	Promega

Table 2.7: Kits and reagents.

## 2.2 Methods

### 2.2.1 General nucleic acid methods

#### 2.2.1.1 Molecular cloning by restriction enzymes digest

##### 2.2.1.1.1 Restriction enzyme digestion

Restriction enzyme (RE) digests were frequently used in the cloning experiments. DNA restriction digest reactions were carried out in a total of 30 µl. As shown below, the mixture consists of 10 units of restriction enzyme (1µl) / 1 µg DNA, 3 µl of the recommended enzyme buffer and an appropriate amount of DNA. For double-digest reactions, only 0.5 µl of each enzyme was used.

Component of a restriction digest reaction	Amount to be added ( $\mu$ l)
10X restriction enzyme buffer	3 $\mu$ l
restriction enzyme (10 Units)	1 $\mu$ l
DNA	X $\mu$ l (up to 1 $\mu$ g)
H <sub>2</sub> O	Up to a final volume of 30 $\mu$ l

The reaction mixture was incubated for a minimum of 1 hour at 37°C. In general, to achieve higher digest efficiency, longer incubation times (up to ~3 hours) were regularly used. REs were heat-inactivated at 65°C for 20 minutes or according to the recommended inactivation conditions provided by the enzyme supplier. Although the RE reactions were more or less identical, a varied amount of DNA were used to obtain the desired DNA yield. These variations will be mentioned in the relevant sections.

#### 2.2.1.1.2 Subcloning

To generate the reporter assay constructs, promoter and ZREs element inserts were ordered as synthetically designed DNA from either Invitrogen GeneArt or GeneStrings services. The GeneArt constructs were provided as a pre-cloned insert in a basic vector. 1  $\mu$ g of this plasmid DNA was double digested with the appropriate enzymes (restriction sites used for each construct are given in the relevant Results chapter). After heat inactivation of the REs, DNA fragments were separated using 1.2% (w/v) agarose gel. The insert DNA band was sliced from the gel under UV light emission. Subsequently, to extract the DNA, the gel slice was processed using gel extraction kit (Qiagen) according to the manufacturer's protocol. The insert DNA concentration was determined based on UV light absorbance using a NanoDrop spectrophotometer (Invitrogen).

Although the promoter constructs ordered using GeneStrings were subcloned in largely the same way as GeneArt constructs, there were a few differences. In contrast to GeneArt constructs, only 200 ng of Invitrogen GeneStrings was used. This was done to minimise the loss of DNA as only a small amount of GeneStrings DNA is supplied (between 400-900 ng). Also, the gel extraction

step was avoided as these constructs were provided as long double-stranded DNA strings. Instead, RE digested inserts were directly purified using the gel extraction kit (Qiagen).

Similarly, the destination plasmid (vector) was digested with same enzymes to obtain compatible sticky ends with the insert. Furthermore, to reduce incompletely-digested vector DNA self-ligating, the vector plasmid was treated with rAPid alkaline phosphatase (Roche). The dephosphorylation step compatibility with the ligation reaction eliminated the need for further DNA purification prior to ligation.

The vector dephosphorylation reaction (described below) was incubated for 10 minutes at 37 °C followed by phosphatase enzyme inactivation at 75°C for 2 minutes.

Component of alkaline phosphatase reaction	Amount to be added (µl)
10X Buffer	2 µl
Enzyme (1 Unit)	1 µl
DNA	X µl (up to 1 µg)
H <sub>2</sub> O	Up to a final volume of 20 µl

#### 2.2.1.1.3 Ligation

An appropriate amount of the insert DNA and digested destination plasmid (vector) was used to set up ligation reaction. Ligation reactions were performed using T4 DNA ligase (Invitrogen). Vector and insert DNA were added in a ratio of 1:3 (vector: insert); typically, 50 ng of vector DNA was used for each ligation while the insert DNA amount was calculated according to the formula below:

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{molar ratio of} = \frac{\text{insert}}{\text{vector}}$$

The final reaction was made in a total volume of 10-15  $\mu\text{l}$  as follows:

Ligation reaction component	Amount to be added ( $\mu\text{l}$ )
10X Ligase Buffer	1 $\mu\text{l}$
T4 Ligase	1 $\mu\text{l}$
Insert DNA	X $\mu\text{l}$
Vector DNA	X $\mu\text{l}$ (50 ng vector)
H <sub>2</sub> O	Up to a final volume of 10-15 $\mu\text{l}$

The ligation samples were incubated in ice/water bath overnight. The following day, a bacterial transformation reaction using 2-3  $\mu\text{l}$  of the ligation mix was set up, as will be explained in the following sections.

### 2.2.1.2 Molecular cloning by In-Fusion ligation

#### 2.2.1.2.1 PCR reaction

In order to generate the His-GFP-bZIP Zta expression construct, In-Fusion ligation approach was used. These experiments were carried out with the help of Dr Louise Bird at Oxford Protein Production Facility (OPPF) as part of a collaboration project. PCR was used to amplify the desired insert from template DNA (His-full length Zta plasmid). Specific PCR primers were designed as described in **Section 2.2.9**.

The lyophilised oligonucleotides were resuspended in nuclease-free water at 100  $\mu\text{M}$  and a working stock for each primer was made at 10  $\mu\text{M}$ . Then, the PCR reaction was set as follows:

PCR reaction component	Amount to be added ( $\mu\text{l}$ )
2X Phusion Flash Master Mix	25 $\mu\text{l}$
Primers (10 $\mu\text{M}$ ) F&R	3 $\mu\text{l}$ each
Template DNA	2 $\mu\text{l}$ (10 ng vector)
ddH <sub>2</sub> O	Up to a final volume of 50 $\mu\text{l}$



The thermal cycling parameters were as the following:

98 °C for 10 seconds

98 °C for 1 second	}	30 cycles
60 °C for 5 seconds		
72°C for 15 seconds		

72°C for 2 minutes

Hold at 4°C

A small amount (5 µl) of the PCR reaction was loaded onto 1.6% agarose gel to ensure a specific PCR product was obtained. Then, the PCR product was treated with DpnI enzyme to digest the template DNA as this DNA could interfere with the subsequent transformation procedure. This is due to the fact that the template DNA (His-Zta) is a plasmid DNA that shares the same antibiotic resistance cassette with the destination vector (pOPINN). DpnI enzyme digests only methylated DNA such as the plasmid DNA replicated within *E. coli*, while synthetically generated DNA (e.g. PCR product) remains undigested. After the DpnI treatment, the PCR product was purified using AMPure XP magnetic bead purification (Beckman Coulter, A63880) according to the manufacturer's protocol.

#### 2.2.1.2.2 In-Fusion Ligation

In-Fusion Ligation Master Mix was ordered as lyophilised mixture. Subsequently, 2 µl of the insert (the PCR purified product) was mixed with a 100 ng of the destination linearised vector (pOPINN). The volume was brought up to 10 µl using nuclease-free water. The mixture was mixed with the ligation lyophilised material and was incubated for 30 minutes at 42 °C in a water bath. After that, the ligation reaction was diluted with 40 µl using nuclease –free water. Only 1 µl of the diluted reaction was transformed into competent bacterial cells. Subsequently, the bacterial cells were plated out on agar plates and processed as explained in more details in the next sections.

### 2.2.1.3 Bacterial transformation

For a typical transformation reaction, a small amount of plasmid DNA (<10 ng) was mixed with 50 µl of competent *E. coli* cells. Strains of *E. coli* used routinely in the various experiments with the appropriate antibiotic concentrations are listed below:

- 1- TOP10 *E. coli* competent cells used for PGL3 and pCDNA3 vectors; 100 µg/ml ampicillin was used for selection.
- 2- PIR *E. coli* competent cells used for pCpGL constructs; 25 µg/ml Zeocin was used for selection).
- 3- Rosetta pLysS cells (protein expression strain) used for pCpGL constructs; 50 µg/ml of Ampicillin and Chloramphenicol (34 µg/ml) were used for selection.

The competent cells aliquot (stored at -80 °C) was allowed to thaw slowly on ice for a few minutes. Then, after the addition of plasmid DNA, the transformation mixture was left on ice for another 20 minutes. To facilitate the uptake of the plasmid DNA by competent cells, the sample was heat-shocked at 42 °C for 45 seconds and placed back on ice for 2 minutes. Soon after, to allow the recovery of the heat-shocked competent cells, 1 ml of LB broth was added under sterile conditions. The sample was left vigorously shaking (225 rpm) in an orbital shaker at 37 °C for an hour. Under sterile conditions, only 10% of the mixture was immediately spread onto an ambient temperature LB agar plate supplemented with the appropriate selective antibiotic. The agar plate was placed upside down in a 37 °C incubator overnight (~16 hours). The following day, the plate was examined for colonies. Given the presence of colonies, the plate was stored at 4 °C until further use within a few days.

An exception to this protocol was the transformation conducted using a ligation reaction. A modified amount of the ligated DNA (2-3 µl) was mixed with 50 µl of competent cells. This was based on the ligation reaction final volume; no more than 20 % of ligation mixture was used. Furthermore, two different volumes

(10% and 90%) of the transformed mixture were plated out onto two agar plates for each transformation.

Another exception was in the case of pOPINN vector cloning where agar plates were prepared with the addition of X-Gal and IPTG. This allows the detection of  $\beta$ -galactosidase activity in order to differentiate colonies that have recombinant plasmids (white colonies) from the ones with inefficiently linearised or non-recombinant plasmids (blue colonies).

To screen for successful ligations, a few single colonies (routinely 2-4 single colonies) were picked using sterile tip, and inoculated, under sterile conditions, in separate tubes containing 5 ml of LB/PowerBroth culture (starter culture) supplemented with the appropriate antibiotic. The inoculated broth samples were kept shaking vigorously at 37 °C for at least 16 hours to allow further growth of the transformed *E. coli*. Samples were stored at 4 °C until further analysis.

#### **2.2.1.4 Subcloning validation**

As described in the previous section, 3 ml of the starter LB broth cultures were used to extract plasmid DNA. The obtained plasmid was assessed for the presence of the insert either via REs double digest or PCR (For the In-Fusion ligated vector) followed by agarose gel electrophoresis. Furthermore, validated samples from the RE digest screen were sent for sequencing analysis (MWG Eurofins, UK). Subsequently, sequencing data was analysed using BLAST tools to validate the ligated DNA by ensuring the presence of the correct insert sequence within the relevant vector plasmid.

#### **2.2.1.5 Bacterial glycerol stock**

For long term storage of bacterial stocks, 800  $\mu$ l of the verified starter culture (as described earlier) was mixed with 200  $\mu$ l of 100% sterile glycerol to achieve a final concentration of 20% glycerol. The mixed samples were stored at -80 °C so they can be used as a source of DNA plasmids.

### **2.2.1.6 Plasmid DNA extraction**

#### **2.2.1.6.1 Mini-prep**

Mini-prep plasmid extraction kit (Qiagen) was mainly used to obtain small amounts of plasmid DNA which can be utilized in either REs screens or DNA sequencing. Using 3 ml of LB broth starter cultures, bacterial cells were pelleted, lysed, and subsequently (the lysate) was passed through a DNA binding silica column. The plasmid DNA extraction and purification were conducted precisely according to the manufacturer's protocol.

#### **2.2.1.6.2 Maxi/Midi-prep**

To obtain a high yield of plasmid DNA, a large-scale kit such as Maxi/Midi preps extraction were used. 500 µl of the validated 5 ml LB broth culture (as described in **Section 2.2.1.3**) was diluted in 500 ml of LB broth containing the appropriate antibiotic and incubated at 37 °C overnight with agitation (225 rpm). The plasmid DNA extraction was carried out according to the manufacturer's protocol (Qiagen).

### **2.2.1.7 Agarose gel electrophoresis**

DNA gel electrophoresis was used to separate and visualise DNA fragments. In most experiments, a concentration ranging from 1-1.4% (w/v) agarose gel was used based on the expected sizes of DNA bands. 50 ml of 1x TBE containing an appropriate amount of agarose powder was brought to boiling temperature to dissolve the agarose completely. The solution was allowed to cool to room temperature before adding 10,000X GelRed DNA stain (Biotium) at a final concentration of 4 µl/ 50 ml of agarose solution. Subsequently, the agarose gel solution was poured into a sealed gel tray fitted with a gel comb. The gel was left to set at room temperature for 20 minutes, then, placed into an electrophoresis tank submerged with 1x TBE buffer. Samples were mixed with blue DNA loading dye (Bioline) and loaded into the wells along with a 1kb DNA hyperladder (Bioline) after removing the comb. Electrophoresis running settings were regularly performed at 100 V for 1 hour. DNA bands visualisation was carried out using LI-COR imaging system (600 channel).

## 2.2.2 General protein methods

### 2.2.2.1 Western blots

#### 2.2.2.1.1 SDS-polyacrylamide gel electrophoresis

To separate proteins and conduct western blots, commercially available precast Bis-Tris SDS-PAGE gels were used (NuPAGE, Invitrogen) along with NuPAGE MOPS SDS running buffer. Total protein lysate or cells pellet was boiled for 10 minutes after the addition of 2X Laemmli protein buffer. Samples were subsequently spun down (13,000 rpm, 5 minutes) to precipitate any debris. An appropriate volume of each sample was loaded into the gel wells along with a pre-stained protein molecular weight marker set (SeeBlue marker, Invitrogen). Gel electrophoresis was carried out in a vertical tank at 200 V for 50 minutes.

#### 2.2.2.1.2 Protein transfer and immunoblotting

After carrying out protein separation by electrophoresis, proteins were transferred to a nitrocellulose membrane. This was performed by placing the gel and membrane between sponges and filter papers (transfer sandwich). The sandwich was placed tightly in a transfer cassette and fitted in a tank filled with transfer buffer (25 mM Tris-HCl pH (8.3), 192 mM glycine, 20% (v/v) methanol). The transfer was conducted at 75V for 90 minutes in a cooled tank.

Then, to check if successful protein transfer was achieved, the nitrocellulose membrane was covered with reversible Ponceau S stain for 2 minutes before being washed under running deionized water.

Following the complete transfer of proteins, the nitrocellulose membrane was placed in a small dark box covered with 10 ml blocking buffer (5% dried milk in PBS-Tween). After 60 minutes of gentle rocking motion at room temperature, the blocking buffer was poured down the sink. Subsequently, the membrane was incubated (for at least 12 hours gently rocking at 4°C) with diluted primary antibodies in blocking buffer at the recommended dilution (A list of the antibodies is given in **Table 2-3**). On the following day, the blotted membrane was washed three times (5 minutes each time) in PBS-Tween at room temperature. By using fluorescent western blot method, it was sufficient to incubate the membrane with fluorescently labelled secondary antibodies

(Odyssey, LI-COR) and visualise the protein bands immediately after three PBS-T washes. The LI-COR imaging system was employed to visualise the protein bands under the 700 and 800 channels. The image capture was allowed for 10 minutes to minimise background signal.

#### **2.2.2.2 SimplyBlue SafeStain**

SimplyBlue SafeStain is an alternative Coomassie stain that provides a fast, sensitive and safe detection of proteins. The SDS-PAGE gel was rinsed (3 times) with around a 100 ml of deionized water to remove SDS and buffer salts as these can interfere with the binding of the dye to proteins. Subsequently, the gel was soaked with around ~20 ml of the stain to cover the gel and left for 1 hour at room temperature with gentle shaking. After incubation, the stain was discarded, and the gel was washed twice with around 100 ml water for 1 hour.

#### **2.2.3 Tissue culture**

##### **2.2.3.1 Cell lines and culture conditions**

Various cell lines were used in the different experiments. B-cells such as the lymphoblastoid cell line 3 (LCL#3) (transformed with B95.8 EBV strain) (Sinclair et al., 1994b) and the EBV-negative Burkitt's lymphoma cells (DG75) (Ben-Bassat et al., 1977) were maintained as suspension cultures. These cells were fed with RPMI 1640 medium (GIBCO, Invitrogen) supplemented with 10% (v/v) foetal calf serum and 100 U/ml of penicillin, 100 µg/ml of streptomycin and 2 mM of L-glutamine (GIBCO, Invitrogen). Cultures were maintained at cells concentration between  $3 \times 10^5$  and  $3 \times 10^6$  /ml and passaged every 2-3 days based on cell density.

293-BZLF1-knockout cells (293-BZLF1-KO) are 293 epithelial cells transfected with BZLF1 mutant EBV genome (Feederle et al., 2000). These cells were maintained as semi-adherent cells in RPMI 1640 media as described above with the addition of Hygromycin at a final concentration of 100 µg /ml for selection. To split these cells, it was enough to incubate them with phosphate buffered saline (PBS) (GIBCO, Invitrogen) for a few minutes to dissociate the cells from the flask.

HeLa, epithelial cells derived from cervical carcinoma (Gey et al., 1952), were maintained as adherent cells in DMEM media supplemented with 10% (v/v) foetal calf serum and 100 U/ml of penicillin, 100 µg/ml of streptomycin and 2 mM of L-glutamine (GIBCO, Invitrogen). Cells were passaged (1 in 5) after reaching 60-70% confluency using trypsinisation.

All cells were kept at 37°C in a humidified incubator containing 5% CO<sub>2</sub> and handled with tissue culture standard aseptic techniques at all time.

For long term storage of cells, a freezing medium was prepared using the cryopreservative agent dimethyl sulfoxide (DMSO) and foetal calf serum (15% DMSO (v/v)). Pelleted cells were resuspended in the freezing medium, divided in cryo-vials and placed in a freezing chamber which contains isopropanol to allow slow freezing at -80°C. After few days, frozen cells were moved from -80°C to a liquid nitrogen storage.

### **2.2.3.2 Transient transfections**

#### **2.2.3.2.1 Adherent cells (Effectene transfection kit)**

293-BZLF1-KO and HeLa cells were transfected using a non-liposomal Effectene kit (Qiagen). The manufacturer protocol was followed with slight changes by reducing the amount of Effectene reagent used to 2.5 µl/ 1 µg DNA. For the luciferase reporter assay, on the day before transfection, HeLa cells were seeded in 6-well plates at 4x10<sup>5</sup> cells/well. The cells were transfected with a total of 1 µg of DNA (500 ng expression vector DNA with 500 ng reporter vector DNA) unless otherwise stated. DNA stock solution at 100 ng/µl was used for all 6-well plates transfections to ensure consistency in transfection protocol. After 48 hours, HeLa cells were scraped off the plate, washed once with PBS and collected into two halves for luciferase and western blots.

On the other hand, 293-BZLF1-KO cells were maintained in large flasks to obtain enough cells that can be used in ChIP experiments. Thus, transfection DNA and reagent volumes were scaled up to 20 times to suit large flask transfections. Cells were harvested and cross-linked 48 hours post-transfection as will be described in **Section 2.2.5**.

#### 2.2.3.2.2 Suspension cells

Electroporation was used to transfect DG75 cells used in luciferase reporter assays. To ensure healthy cell growth and viability, cells were split 1:3 on the day before transfection. For each transfection, around  $1.2 \times 10^7$  cells suspended in 300  $\mu$ l, pre-warmed serum-free medium was electroporated. A total of 10  $\mu$ g of DNA, which contains 5  $\mu$ g DNA of either pcDNA3 (control) or Zta expression vector (HisZta-pcDNA3) with 5  $\mu$ g of the luciferase reporter vector was used.

Electroporation was conducted in disposable transfection polycarbonate cuvettes with 4 mm gap size (VWR) using BIORAD Genepulser II. Electroporation conditions were set at 250 voltage and 975 capacitance. Transfected cells were immediately resuspended in 1 ml of media and transferred in 9 ml of media in a small flask. Transfected cells were incubated for 24 hours at 37°C before harvesting. Cells collected (24 hours post-transfection) into two halves: luciferase assay and western blot.

### 2.2.4 Luciferase reporter assays

#### 2.2.4.1 Luciferase reaction

The firefly luciferase reporter assay (Promega) was used to investigate promoters and ZREs constructs activity. Transfected cells (either HeLa or DG75 cells as described in **Section 2.2.3.2**) were pelleted and resuspended in 250  $\mu$ l 1X Passive Lysis Buffer (Promega). To ensure efficient lysis and homogenisation, samples were vortexed for 10 seconds. The cell lysis was allowed to take place at room temperature for 20 minutes which was followed by centrifugation in a microfuge at 13,000 rpm for 5 minutes at 4°C. For each transfection, 10  $\mu$ l of the lysate was loaded in triplicate into a white 96 well plate. The plate was placed in plate reader machine (Promega GloMax multi-detection system) which allows the auto-dispensing of luciferase reagent (Promega LAR reagent). The plate reader was set to dispense 50  $\mu$ l of LAR, take a read for 10 seconds with a delay of 2 seconds after the addition of LAR reagent. Light emission was measured in relative light units (RLU). A background value representing the average of triplicate reads of the Passive Lysis Buffer was subtracted from all reads to account for background signal.



#### **2.2.4.2 Normalization**

As previously stated, harvested cells for luciferase assay were divided equally into two halves. One-half was used to run a western blot to show Zta transfection levels as well as a loading control ( $\beta$ -actin).

Thus, cells were boiled in 125  $\mu$ l sample buffer Laemmli 2X (Sigma) with only 5  $\mu$ l of each sample was loaded onto SDS-PAGE gel. From the western blot, the actin band intensity was quantified for each sample using the LI-COR Image Studio software. For each sample, the individual luciferase RLU reads were divided by the actin band intensity value for that sample. Then, the average and standard deviation were determined for the normalized RLU reads.

#### **2.2.5 Chromatin immunoprecipitation (ChIP-qPCR)**

##### **2.2.5.1 Chromatin preparation**

Zta binding to various sites in the viral genome was determined in B lymphocytes and epithelial cells (LCL#3 and 293-BZLF1-KO cell lines) using ChIP-qPCR. First, cells were pelleted and resuspended using normal media to be at  $1 \times 10^7$  cells/ml. To cross-link DNA-binding proteins to their sites, formaldehyde was added at a final concentration of 1% (v/v), and samples were kept rocking at room temperature for 10 minutes. The cross-linking reaction was stopped by the addition of glycine at a final concentration of 0.125 M.

The cells were pelleted by centrifugation at 1300 rpm for 5 minutes and washed twice with PBS. Pelleted and washed cells were either snap frozen or subsequently resuspended in 600  $\mu$ l per  $1 \times 10^7$  cell lysis buffer (85mM KCl, 0.5% NP-40, and 5mM PIPES pH 8.0). Cell lysis took place while samples were incubated on ice for 10 minutes. Soon after, nuclei were pelleted at 3500 rpm at 4°C for 10 minutes. The pelleted nuclei were resuspended in 200  $\mu$ l ChIP SDS lysis buffer per  $1 \times 10^7$  cells (1% SDS (v/v), 10 mM EDTA, 50 mM Tris pH 8.0) after the removal of the supernatant. At this stage, the chromatin was ready for physical shearing by sonication. Chromatin fragments were reduced to small fragments between 200 and 600 base pairs in length. Sonication was carried out in 1.5 ml Eppendorf tubes in a total volume of 200  $\mu$ l chromatin aliquots.

During the sonication, samples were kept in ice water bath to avoid any effect the excessive heat generated by the sonication probe. The sonication conditions were set for 12 cycles of 10 second on/off pulses at 30% output using Sonics Vibra cell VC 750 sonicator.

#### **2.2.5.2 Immunoprecipitation**

To prepare the beads, 100 µl of mixed protein A and protein G beads of 50% (w/v) bead slurry was used per immune-precipitation (IP). Beads were spun down at 3000 rpm at 4 °C for 5 minutes and washed 3 times with ChIP dilution buffer (0.01% (v/v) SDS, 1.1% (v/v) Triton X-100, 1.2 mM EDTA, 16.7 mM Tris pH 8.0, 167 mM NaCl). After that, the beads were resuspended in ChIP dilution buffer at 55 µl per IP. The beads were then blocked by incubation with 55 µl of 10 mg/ml salmon sperm DNA (Invitrogen), at 4°C and kept rotating for 30 minutes. Following that, the beads were washed twice with IP dilution buffer and resuspended in 55 µl IP dilution buffer.

The chromatin samples were thawed on ice and then diluted 1:10 with IP dilution buffer to a final volume of 1 ml. Chromatin was pre-cleared by adding 45 µl of bead slurry and incubated at 4°C with rotation for 30 minutes. The samples were centrifuged at 3000 rpm at 4°C for 5 minutes, and the supernatant was transferred to new tubes.

At this point, 40 µl was removed from each sample (represents 4% input sample). Then, 10 µg of antibody to Zta (goat polyclonal antibody, sc-17503, Santa Cruz Biotechnology) or control antibody (normal goat IgG, Santa Cruz Biotech sc-2028) were added to each sample and incubated at 4°C with rotation for at least 1 hour. 55 µl of beads were added per IP and incubated at 4°C with rotation overnight.

The following day, the beads were recovered by centrifugation at 3000 rpm at 4°C for 5 minutes. The beads were then washed once (unless stated otherwise) at different contingency conditions using the following buffers in order:

- 1- Low salt buffer (0.1% (v/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.0, 150 mM NaCl),

- 2- High salt buffer (0.1% (v/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.0, 500 mM NaCl)
- 3- LiCl buffer (250 mM LiCl, 1% (v/v) NP-40, 1% (v/v) Na-deoxycholate, 1 mM EDTA, 10 mM Tris pH 8.0)
- 4- ChIP TE buffer (0.01M Tris-HCl 5mM EDTA pH 7.5) (two washes).

Each wash was for 15 minutes at 4°C while samples are rotating ChIP. Also, antibodies complex bound to beads were eluted with ChIP elution buffer (0.01M Tris-HCl, 5mM EDTA pH 7.5, 1% (v/v) SDS) by incubation at 65°C for 15 minutes. At this step, input samples were thawed and processed in the same way as the IP samples. The supernatant was recovered and was incubated at 65°C overnight to reverse cross-linking.

The following day, the samples were diluted 1:2 with 150 µl of ChIP TE buffer and 2.5 µl of proteinase K (20 µg/µl) was added. To ensure the removal of all the protein, the samples were incubated at 55°C for at least 2-3 hours. Finally, the DNA was purified using PCR purification kit (Qiagen) and eluted in 100 µl nuclease-free water.

### 2.2.5.3 Real time-qPCR

Absolute quantification qPCR was performed using GoTag (Promega). The reaction was set as follows:

2X Promega Master Mix	12.5 µl
Primers Forward and Reverse mix (5 µM each)	1.25 µl
Nuclease-free water	8.75 µl
DNA template	2.5 µl

The samples were run in 96-well plates and analysed using Applied Biosystems StepOne real-time PCR machine. The standard curves were obtained by making up 5 serial dilutions starting with 4% of the input DNA sample. The cycle's parameters were as the following:

95°C for 10 minutes

40 cycles of:

95°C for 15 seconds

60°C for 1 minute

Primers used in the ChIP-qPCR experiments for target regions are listed in **Table 2.4**.

### **2.2.6 Protein expression and purification (His-GFP-bZIP Zta)**

To investigate Zta interaction with BNLF2a ZREs, a basic leucine zipper domain of Zta tagged with His-GFP vector was cloned and screened for protein expression, as previously described.

This vector was used to express bZIP Zta domain (tagged with GFP and poly histidine) in *E. coli* at large scale. Then it was purified using histidine affinity column as well as a size exclusion chromatography column.

#### **2.2.6.1 Protein expression**

A single colony of transformed Rosetta PlysS competent cells was inoculated into 20 ml of Powerbroth supplemented with the appropriate antibiotics in a 50 ml tube. After overnight incubation at 37°C, with rigours shaking at 225 rpm, 10 ml of the overnight culture was added to 500 ml of Overnight Express Instant TB Medium (autoinduction media) in 2 L flask supplemented with the appropriate antibiotics. The OD was continuously monitored while the culture was incubated at 37°C and shaking at 225 rpm. The incubation temperature was reduced to 25°C when the OD reached around 0.6-0.8 and the culture was incubated for approximately a further 20 hours.

### 2.2.6.2 Protein purification

Transformed Rosetta PlysS cells from the autoinduced overnight culture were harvested at 6,000 g for 15 minutes at 4°C. The pellet was weighed and resuspended in an appropriate amount of lysis buffer (based on the weight of the pellet (30 ml/10 g of pellet)). This was followed by sonication using 10 seconds off/on pulse for 10 minutes. The lysate was spun again to remove the insoluble fraction at 30,000 g for 30 minutes at 4°C. The clear lysate was transferred to a clean tube and kept on ice.

A 1 ml HiTrap TALON Cobalt2+ column (GE Healthcare) was connected to FLPC ÄKTA purifier system and washed according to the manufacturer's recommendations. Then, equilibrated with 15 ml of equilibration buffer (50 mM Tris-CL pH 7.5, 500 mM NaCl, TCEP 0.5 mM). 50 ml of the lysate was injected into the super loop (connected to the ÄKTA system). The machine was set up to inject the lysate through the column at 1 ml/min flow, perform a 10 ml column wash using washing buffer, and elute using 10 ml elution buffer by collecting 20 fractions (0.5 ml each) in a 96 deep well-plate. The fractions under the highest UV trace peak were pooled together and concentrated to 1 ml using VIVASPIN MWCO 50K spin column.

On the same day, a GF 75 size exclusion chromatography column (GE Healthcare) was connected to FLPC ÄKTA purifier system and prepared according to the manufacturer's recommendations. Subsequently, 500 µl of the concentrated protein sample (post affinity column purification) was injected into the column. The protein molecules were allowed to separate based on size by collection 60 fractions (1 ml each) in elution buffer (50 mM Tris-HCL, 300 mM NaCl, TCEP 0.5 mM). SimplyBlue stained SDS-PAGE was used to inspect the presence of the expected size band of the purified protein in the collected fractions in correlation with the UV trace.

### 2.2.7 HeLa cell nuclear extract

Around  $2 \times 10^7$  cells were scraped off the flask in 10 ml ice-cold PBS, pelleted and washed in 5 ml cold PBS. Pelleted cells were resuspended in 200  $\mu$ l cytoplasmic extract buffer and kept on ice for 3 minutes to allow the breakage of the cells. Then, intact nuclei were pelleted in a microfuge at 1500 rpm for 4 minutes at 4°C. The pelleted nuclei were washed once in 200  $\mu$ l cytoplasmic extract buffer without NP-40. A 100  $\mu$ l of nuclear extract buffer was added to the pelleted nuclei and kept on ice for 10 minutes with interval vortexing every 2-3 minutes. The extract was aliquoted into smaller volumes, snap frozen, and stored at -80°C.

### 2.2.8 Electrophoretic mobility shift assay

EMSA assays were performed to investigate proteins of interest interactions with DNA *in vitro*. This requires the mixture of labelled DNA probes with the protein of interest followed by gel electrophoresis.

#### 2.2.8.1 Fluorescent EMSA probes preparation

The EMSA probes were ordered as 27 bases single-stranded complementary oligonucleotides from Integrated DNA Technologies (IDT). Only one strand of the complementary oligos was labelled at the 5' end with an IRDye 800 dye. Lyophilised labelled, and non-labelled oligos were resuspended in nuclease-free water to make a 100  $\mu$ M stock of each oligo. A working stock at 20  $\mu$ M (20 pmol/ $\mu$ l) was made by diluting each oligo in 1X EMSA annealing buffer. Then, equal volumes were mixed of the labelled complementary oligo and a 3-fold molar in excess of the unlabelled strand. To allow annealing of both strands, samples were placed in a 100°C heating block for 5 minutes. Subsequently, the heating block was turned off and samples were allowed to cool slowly until reaching room temperature. An EMSA working solution of the annealed DNA strands (EMSA probe) was achieved by diluting the annealed oligonucleotides at 1 in 10 ratio using nuclease-free water. EMSA probes were stored protected from light at -20 °C until further use. A list of all EMSA oligonucleotide sequences is provided in **Table 2-5**.

### 2.2.8.2 EMSA binding reaction and gel electrophoresis

The EMSA binding reaction was carried out as listed below:

Binding reaction	Amount
5X binding buffer (Promega)	4 $\mu$ l
DNA probe (EMSA working stock)	2 $\mu$ l
Protein:	
- HeLa nuclear extract (5-30 $\mu$ g)	X $\mu$ l
- Zta Purified protein (600 ng)	
DTT (0.1M)	0.8 $\mu$ l
Nuclease -free water	X $\mu$ l
Total:	20 $\mu$ l

Then, the EMSA binding reaction was allowed to take place for 30 minutes on ice/water bath while being protected from light. For the super shift EMSA, HeLa nuclear extract was incubated with 2  $\mu$ g of the relevant antibody for 20 minutes prior to the addition of the EMSA reaction mixture.

Meanwhile, a pre-run of the pre-cast native DNA retardation gel, which consists of 6% (v/v) polyacrylamide in 0.5XTBE buffer, was carried out in a pre-chilled 0.5 X TBE buffer at a constant voltage (e.g. 100 V) for around 30-45 minutes.

After the incubation, the binding reaction was mixed with 2  $\mu$ l of loading dye (10X orange dye, LI-COR). 10  $\mu$ l of each sample was loaded onto the gel and was allowed to run for 1 hour at 100 V. Once the run is done, the gel was imaged in the LI-COR under the 800 channel allowing 10 minutes for image acquisition.

## 2.2.9 Computational methods

### 2.2.9.1 Primers design

New primers were designed to investigate Zta binding at specific viral DNA regions where previous primers were not available in our lab. These oligos were designed using NCBI Primer-Blast web tool ([www.ncbi.nlm.nih.gov/tools/primer-blast](http://www.ncbi.nlm.nih.gov/tools/primer-blast)). The oligos were created to be around 20 bp and give a product length ranging from 90 to 150 bp. Primers specificity was checked against the EBV wildtype genome (NC\_007605.1) and EBV B95-8 (V01555.2) as well as the human genome. Design parameters were adjusted to obtain approximately 65°C T<sub>m</sub> and 55% and 65% GC content for each oligo. On the other hand, primers used in the In-Fusion ligation were designed by Dr Louise Bird at Oxford Protein Production Facility (OPPF) using their automated primer design tool, which is available at <https://www.oppf.rc-harwell.ac.uk/Opiner>.

### 2.2.9.2 MEME and TOMTOM motif search

The conserved ZREs flanking motif was identified *in silico* by entering each BNLF2a core ZRE site (7 bp) and 14 bp flank bases on each side into the MEME motif search web tool (MEME Suite 4.10.2 <http://meme-suite.org>) (Bailey et al., 2009b). The significant transcription factor motif that were discovered using MEME was then compared with other transcription factor motifs using the motif database scanning algorithm TOMTOM (Bailey et al., 2009b).

### 2.2.9.3 ALGGEN-PROMO transcription factor motifs predicting tool

PROMO web tool was used to search for Zta binding sites (ZREs). This tool predicts various transcription factors binding sites (TFBS) based on weight matrices defined in the TRANSFAC database (version 8.3) (Farre et al., 2003). The web tool is accessible online on: [http://alggen.lsi.upc.es/cgi-bin/promo\\_v3/promo/promoinit.cgi?dirDB=TF\\_8.3](http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3)

### 2.2.9.4 Diagram and illustration design

Thesis diagrams were designed using Adobe Illustrator CS6 and Microsoft PowerPoint.



## Chapter 3 Zta binding across the viral genome

### 3.1 Introduction

As discussed in **Chapter 1**, upon EBV entry into the cell, the circularised viral DNA is maintained as an episome packed in higher order repressive chromatin structure in a very similar way to the host genome. This chromatin state is essential to maintaining latency by limiting the viral expression program to a specific set of genes (reviewed in (Lieberman, 2013)).

However, the viral transactivator Zta is capable of disrupting this latency (Countryman and Miller, 1985; Grogan et al., 1987). This event takes place in spite of the strong association between the viral genome and repressive epigenetic marks. Thus, Zta must have evolved with unique characteristics that allow it to overcome these repressive marks to induce lytic cycle. In fact, Zta plays a vital role in orchestrating many of the early events that take place during lytic cycle including the regulation of host and viral genes as well as viral DNA replication. These events are achieved by the ability of Zta to bind directly to a wide range of binding motifs ( AP-1 and ZREs sites) (Farrell et al., 1989; Urier et al., 1989) in methylated and/or unmethylated states (Bhende et al., 2004, 2005; Wille et al., 2013). Also, Zta interacts with the cellular general transcription factors such as TFIID and TFIIA (Lieberman, 1994; Lieberman and Berk, 1991; Lieberman et al., 1989; Lieberman et al., 1997). More importantly, Zta is known to recruit histone-modifier proteins which lead to a more accessible chromatin. For example, Zta is known to directly bind histone acetylase CBP (Adamson and Kenney, 1999; Deng, 2003; Zerby et al., 1999). Also, Zta is known to associate with the repressive histone mark (H3K9me3) across various early lytic cycle promoters (Ramasubramanyan et al., 2012b).

These multifunctional features of Zta have generated much interest in investigating its structure and function in the past few years. In line with this, the powerful technique of chromatin immunoprecipitation (ChIP), coupled with second generation DNA sequencing (ChIP-Seq) was employed to investigate Zta binding across the viral and cellular genome (Bergbauer et al., 2010;

Ramasubramanyan et al., 2012a; Ramasubramanyan et al., 2015). The previous studies have established Zta binding patterns in B cells only.

On the other hand, EBV infects and replicates in epithelial cells (Greenspan et al., 1985; Sixbey et al., 1984). EBV latency is not supported in normal terminally differentiated epithelial cells (Nawandar et al., 2015; Niedobitek et al., 1991; Young et al., 1991c). This could be attributed to Zta continuous expression in differentiated epithelial cells (Young et al., 1991c). What is more interesting is that epithelial cells growing in organotypic cultures showed a high rate of EBV replication (Temple et al., 2014). To understand the features of EBV lytic cycle in epithelial cells, genome-wide Zta binding analysis will be critical. Although there is no global data for Zta binding in an EBV epithelial system published to date, a genome-wide analysis using ChIP-Seq for Zta binding was carried out recently in an EBV-positive NPC cell line by Anja Godfrey, a Ph.D. student in our lab. This will provide an insight into EBV lytic cycle differences between B cells and epithelial cells.

Here, we will focus on Zta binding across the viral genome only. In particular, differences observed in Zta binding in different EBV-positive cell lines in the previously published ChIP-Seq data sets as will be discussed in the next sections. Using ChIP coupled with qPCR (ChIP-qPCR), we extend Zta binding analysis in two different EBV systems including HEK 293 cells infected with EBV genome cloned on a bacterial artificial chromosome (BAC) (Feederle et al., 2000), a system that has been extensively used in EBV genome mutational analysis (Feederle et al., 2010). Furthermore, we address conservation of Zta enrichment in different cell lines at novel EBV promoters and genomic regions previously described for the first time (Ramasubramanyan et al., 2012a).

## 3.2 Results

### 3.2.1 ChIP-qPCR approach to analyse Zta binding to EBV genome

Our specific aim was to investigate the differences between various Zta binding patterns seen in published ChIP-Seq data as described in **Section 3.1**. An example of these differences is given in **Figure 3-1 (A and B)** where Zta binding to the EBV genome was compared between different cell lines analysed in two separate published ChIP-Seq data sets (Bergbauer et al., 2010; Ramasubramanian et al., 2012a). Zta binding peaks across the viral genome from the two sets of data (aligned according to B95-8 EBV genome (V01555)) were aligned and regions of differences identified.

Two regions that illustrate the major differences in Zta binding are shown in **Figure 3-1 (A and B)**. First, there was an apparent Zta binding peak between 0-2 kbp in EBV-positive Akata BL cells previously identified by members of our lab (Ramasubramanian et al., 2012a) but not in the other two cell lines (Raji and B95-8 cells)(Bergbauer et al., 2010). This binding peak occurs in the region of the EBV late lytic *BNRF1* promoter.

Similarly, in **panel B**, a Zta binding peak at ~95 kbp (according to B95-8 EBV genome) is seen in Akata BL cells but not in the other two cell lines. A query of this site in NCBI genome browser showed that the region is not associated with any regulatory function. It is worth to note that this region aligns to a region around 82 kbp in the wildtype EBV strain (NC\_007605). This region will be discussed further and referred to as ~82 kbp in the following sections.

Thus, there is strong evidence that Zta binds differently across the viral genome in different EBV systems derived from B cells despite the strong correlation of Zta binding observed at most of the regions across the viral genome.

To investigate the differences seen in Zta binding patterns across the viral genome, a ChIP-qPCR approach was taken. Additionally, we expanded this investigation of Zta binding in an EBV non-B cell system. Thus, two different cell lines, LCL#3 which are lymphoblastoid B-lymphocytes in latency III (Sinclair et al., 1994b) and 293-BZLF1-KO cells, a model for EBV non-B cell system

(Feederle et al., 2000), were used, as will be seen in later experiments. In LCL#3, a small sub-population of these cells continuously goes into lytic phase. On the other hand, 293-BZLF1-KO cells, which were made by infecting HEK 293 cells with recombinant EBV DNA with a deletion in *BZLF1* gene, require Zta ectopic expression to induce lytic cycle.

The ChIP-qPCR experimental design is illustrated in **Figure 3-2**. LCL#3 and 293-BZLF1-KO cells were maintained, as described in **Chapter 2**. In the case of 293-BZLF1-KO cells, His-full length Zta plasmid (His-Zta) was transfected 48 hours prior to harvesting to induce lytic cycle. Zta expression was verified using western blotting, as shown in **Figure 3-2 (A)**.

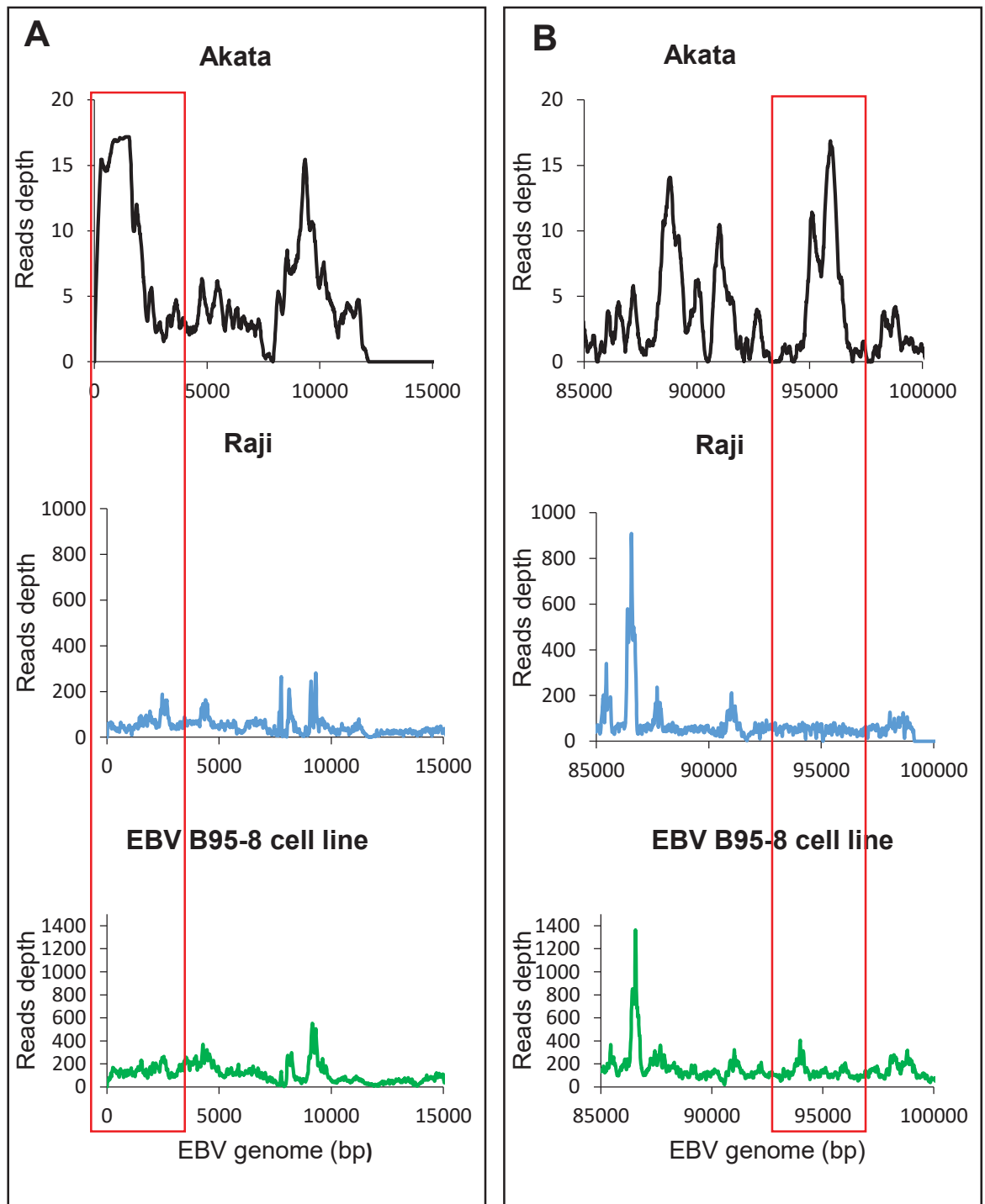
Each cell line was harvested and treated with formaldehyde to cross-link proteins to DNA. 293-BZLF1-KO cells were harvested 48 hours post-transfection. Upon lysis, the cross-linked chromatin was sonicated to shear it to small fragments. The sonicated material was reversed cross-linked and a sample was checked on 1% agarose gel to assess the shearing efficiency as shown in **Figure 3-2 (B)**. An enriched DNA smear between 200-600 bp indicated successful chromatin shearing.

Chromatin immunoprecipitation was performed using protein A and G Sepharose beads along with a specific antibody against Zta and a control antibody (raised in the same species) in LCL#3 experiment. On the other hand, for 293-BZLF1-KO cells, two chromatin preparation were used; one transfected with His-Zta plasmid, and a control sample transfected with pcDNA3 plasmid as shown in the western blot (**Figure 3-2 (A)**).

After the immunoprecipitation, samples were reversed cross-linked, and treated with proteinase K to remove the protein in the samples. This was followed by DNA purification and qPCR to determine whether enrichment for the targets of interest had occurred. The primers used for qPCR are listed in a **Table 2-4** in **Chapter 2**.

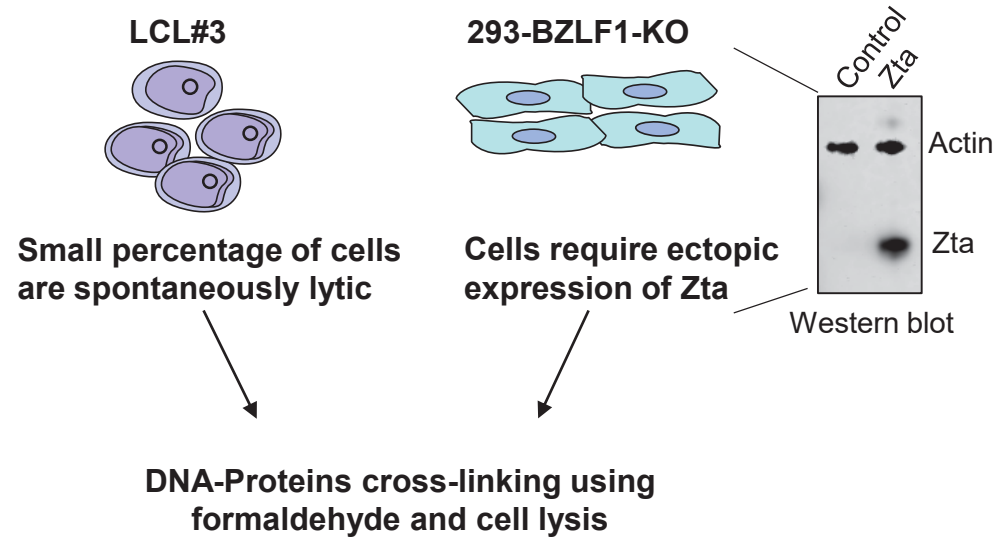
The amplification efficiency and specificity of the primers was checked by adding a melting curve step in the qPCR protocol. The primers had a range of 80% and 98% amplification efficiency. In addition, the standard curves used to measure the absolute quantity of DNA had slope values between -3.755 to -3.365 and  $R^2$  values around 0.998. This indicates consistency in the amplification of the serial dilution samples seen in the standard curves as illustrated in the example provided of the melting and standard curves for one of the primers (**Figure 3-2 (C)**).

Overall, a well-established ChIP-qPCR protocol and previously described ChIP-grade antibody against Zta were used in these experiments. In addition, various quality control steps as illustrated in **Figure 3-2** were carried out at every step to minimise any technical variations.

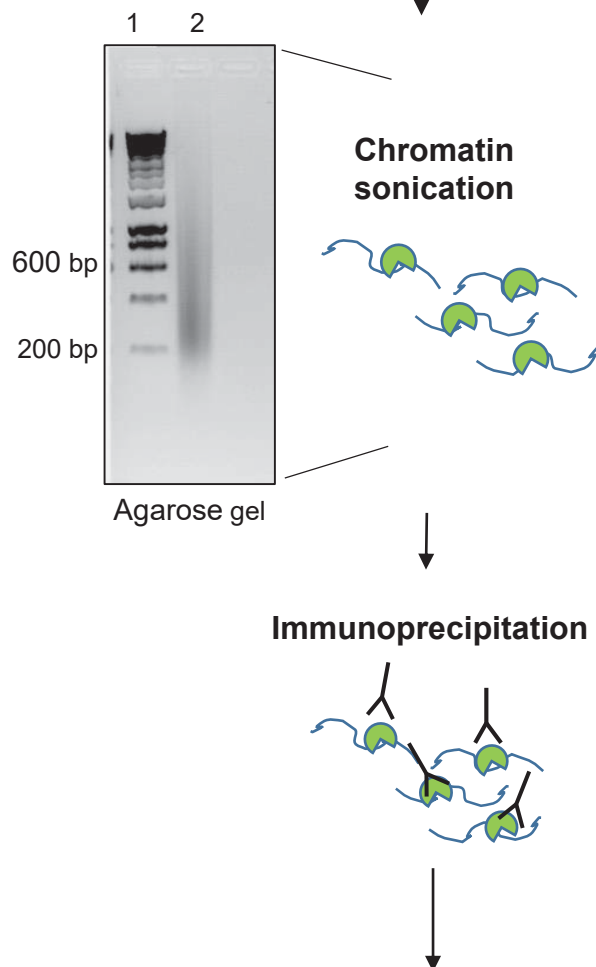


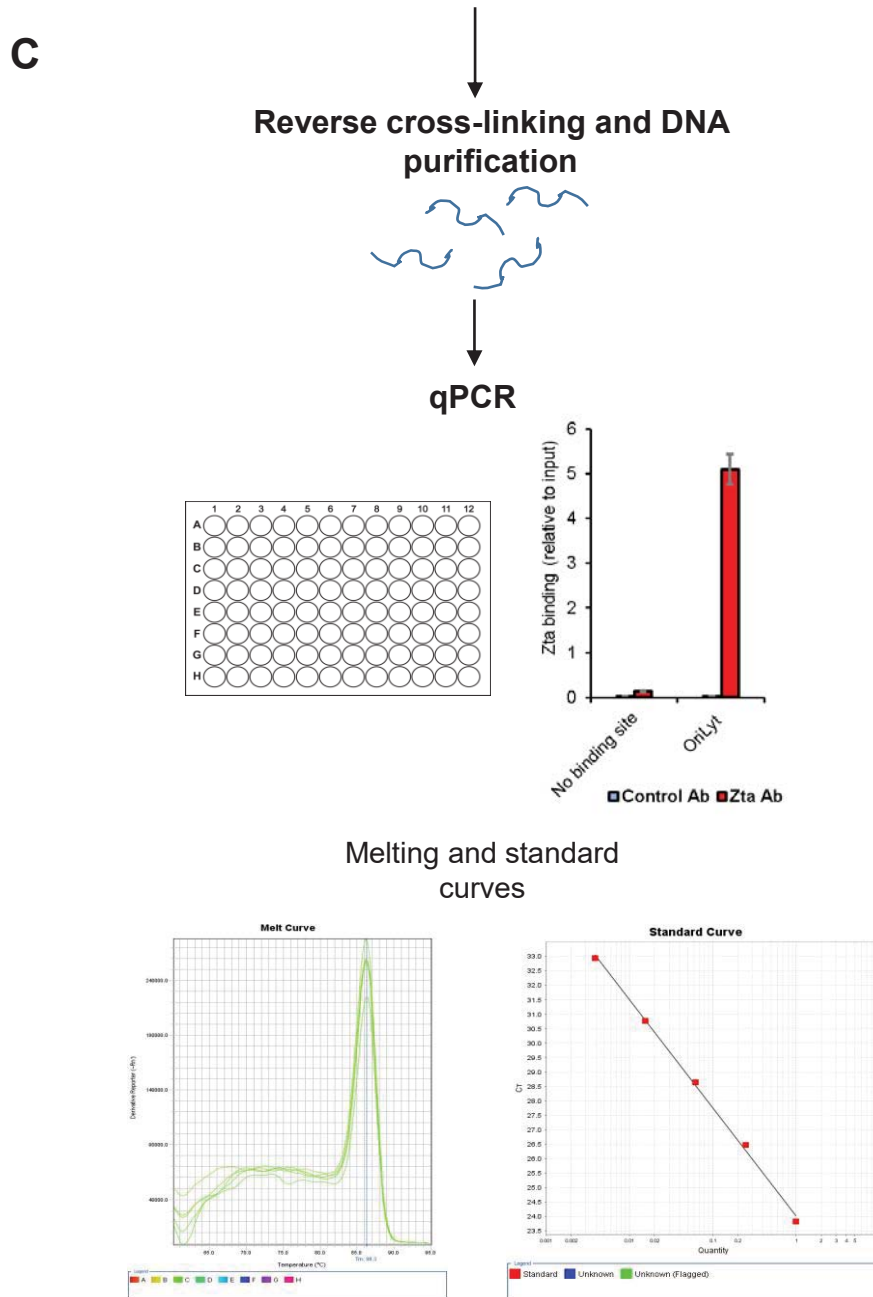
**Figure 3-1 Differences in Zta binding to the EBV genome.** Zta binding to the EBV genome was compared between two different published ChIP-Seq data sets. Zta binding peaks were aligned according to B95-8 EBV genome (V01555) in EBV-positive Akata (black peaks), Raji (blue peaks) and B95-8 (green peaks) cell lines. **(A)** Binding at 0-2 kbp of EBV genome was seen in Akata but not in the other two cell lines (red box). **(B)** Similarly, binding between 95-97 kbp of EBV genome (red box) was seen in Akata but not in the other two cell lines. This region aligns to a region around 82 kbp in the wildtype EBV strain (NC\_007605). It will be referred to as ~82 kbp in the following sections. ChIP-seq data were taken from: 1- Akata cells (Ramasubramanyan et al., 2012). 2- Raji and B95.8 cells (Bergbauer et al., 2010).

**A**



**B**





**Figure 3-2 Experimental design of Zta ChIP-qPCR.** (A) Zta binding patterns were analysed in two EBV systems: the B cells LCL#3, which are spontaneously lytic and the non-B cells 293-BZLF1-KO, which require Zta ectopic expression to induce lytic cycle (western blot). Cells were harvested (48 hours post-transfection for 293-BZLF1-KO) and treated with formaldehyde to cross-link proteins to DNA. (B) After cell lysis, chromatin was sonicated to produce short fragments between 200-600 bp (as seen on the agarose gel (Lane 2)). DNA bound to Zta was precipitated using specific Zta antibody. (C) Finally, qPCR was used to quantify enriched fragments using specific primers for the target regions. Along with Zta antibody, a control antibody was used in LCL#3 ChIP. On the other hand, in 293-BZLF1-KO experiment, a control sample (293-BZLF1-KO transfected with control plasmid) treated identically to Zta sample was used. A serial dilution of the input DNA was prepared to generate a standard curve which was used to quantify enriched DNA fragments. qPCR melting and standard curves were used as an indicator of product specificity and amplification efficiency, respectively.

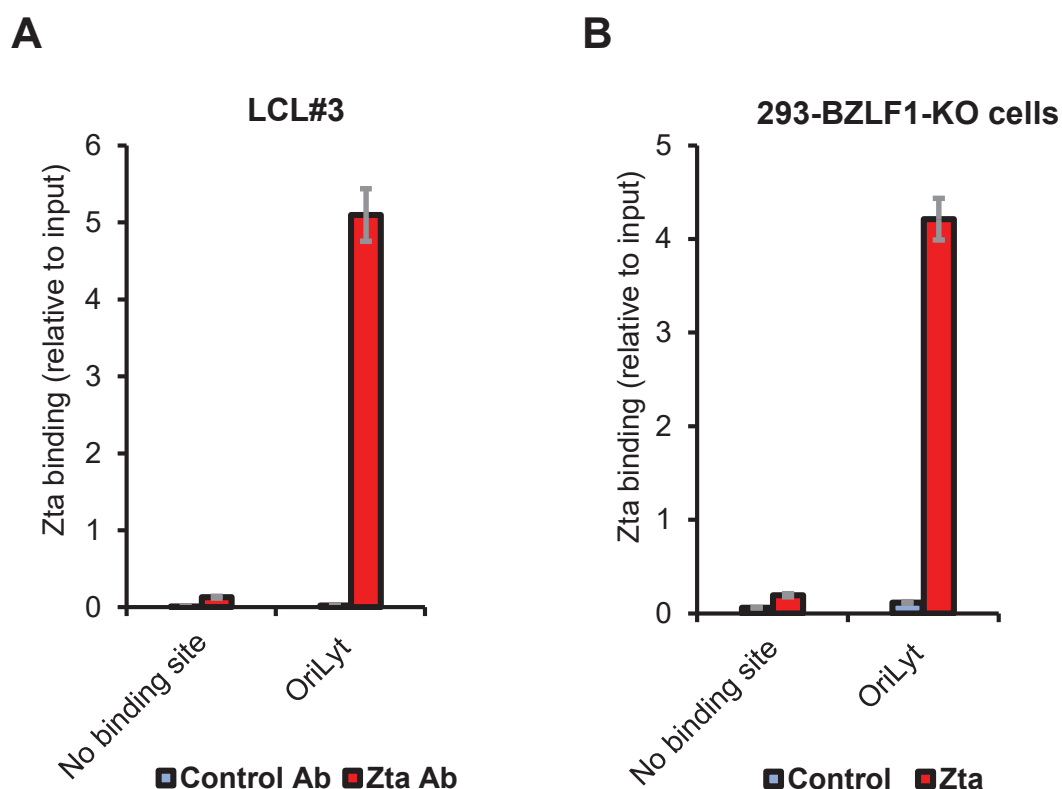


### 3.2.2 Zta binds similarly to EBV OriLyt

To address Zta association with EBV origin of lytic replication (OriLyt) in the two cell lines, ChIP-qPCR specific primers for this target region as well as a flanking region (negative control) were used. We looked into the OriLyt region because the association between Zta and OriLyt is well-established and is essential for viral DNA replication. The EBV genome contains two identical OriLyt sites. These sites are functionally redundant, as EBV B95-8 strain still goes into full lytic cycle despite a deletion spanning one of the OriLyt regions. The primers used in the qPCR are specific to both OriLyt sites.

The qPCR data presented in **Figure 3-3** showed significant enrichment for Zta at OriLyt compared to OriLyt-Flank region for both LCL#3 and 293-BZLF1-KO cell lines (**Figure 3-3 (A and B)**). Also, this enrichment is significant compared to the control samples in both cell lines. Zta binding at OriLyt is higher than other sites, as will be seen in the following figures.

Our finding indicates that there was no significant variation in Zta association with EBV OriLyt during lytic cycle between the two EBV-positive cell lines.



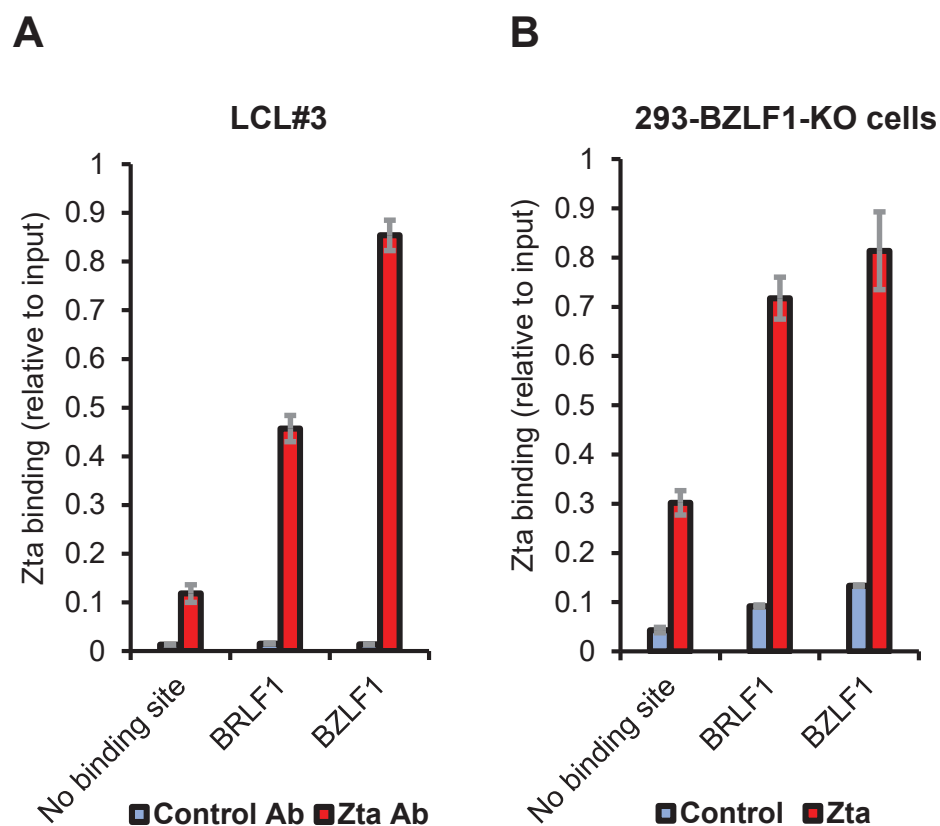
**Figure 3-3 Zta binding at the origin of lytic replication (OriLyt).** ChIP-qPCR in different EBV cell lines: **(A)** LCL#3 cells and **(B)** 293-BZLF1-KO. Zta binding (red bars) is compared to a control sample (blue bars). The qPCR assays were performed using primers (provided in Table 2-4) for the origin of lytic replication (OriLyt), and a negative control region flanking OriLyt (no binding site). Values represent absolute levels of enrichment calculated using a standard curve of the input material. Error bars represent the mean of triplicate readings  $\pm$  SD for each sample.

### 3.2.3 Zta binds similarly to promoter regions of the immediate-early genes

The association between Zta and promoter regions of the immediate-early genes (*BZLF1* and *BRLF1*) was assessed using ChIP-qPCR in two cell lines, as previously described. Zta is known to regulate *BRLF1* and its promoter during very early lytic events (Sinclair et al., 1991). Also, Zta autoregulates its own promoter through two ZRE elements (ZIIIA and ZIIIB) (Flemington and Speck, 1990a). Specific primers to these target regions as well as the negative control OriLyt-Flank region were used (**Table 2-4**). 293-BZLF1-KO cell line contains a deletion of *BZLF1* gene that affects only part of exons 1 and 2 but not the promoter site (Feederle et al., 2000).

The qPCR data showed significant enrichment for Zta at *BZLF1* and *BRLF1* promoters compared to the OriLyt-Flank control region (no binding site) in both LCL#3 and 293-BZLF1-KO cell lines (**Figure 3-4 A and B, respectively**). This enrichment is not as pronounced as what was seen at OriLyt region, but similar findings were seen in early and late lytic promoters (**Figure 3-5 and Figure 3-6**).

Our findings indicate that there was no significant variation in Zta association with EBV immediate-early lytic promoters during lytic cycle between the two EBV-positive cell lines used in this experiment.



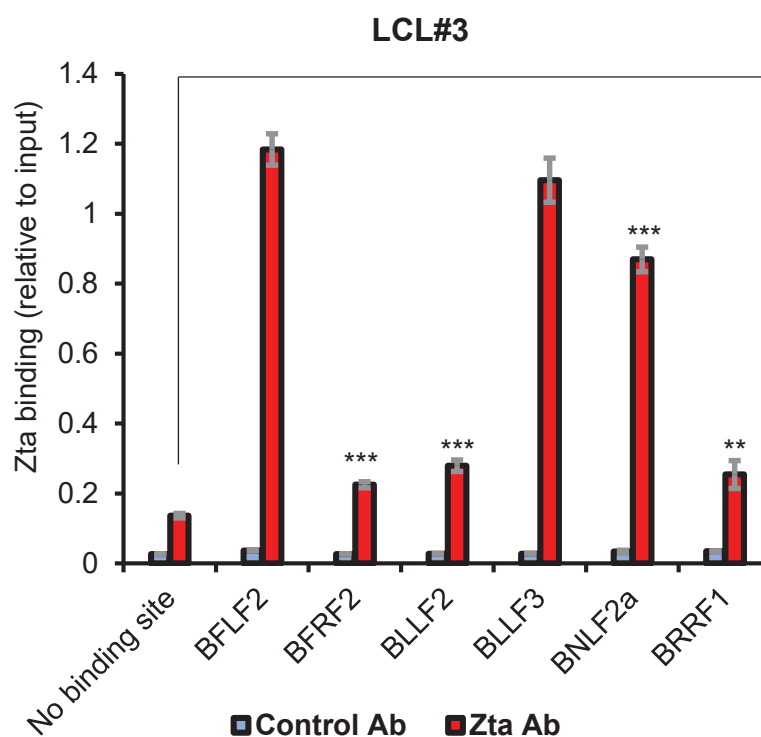
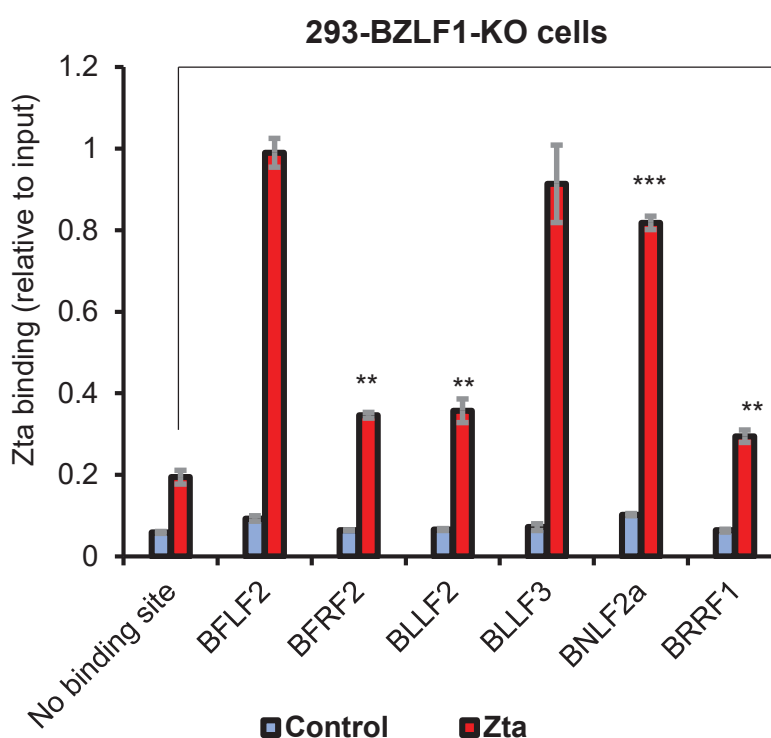
**Figure 3-4 Zta binding at the immediate-early lytic promoters.** ChIP-qPCR in different EBV cell lines: **(A)** LCL#3 cells and **(B)** 293-BZLF1-KO. Zta binding (red bars) is compared to a control sample (blue bars). The qPCR assays were performed using primers (provided in Table 2-4) for no binding site (OriLyt-Flank) and the immediate-early promoters of *BRLF1* & *BZLF1* genes. Values represent absolute levels of enrichment calculated using a standard curve of the input material. Error bars represent the mean of triplicate readings  $\pm$  SD for each sample.

### 3.2.4 Zta binds similarly to promoter regions of various early lytic genes

We further assessed the association between Zta and various promoters of EBV early lytic genes. This was of great interest because the association between Zta and some of these targets were first described only recently (Bergbauer et al., 2010; Ramasubramanian et al., 2012a).

The qPCR results showed a significant enrichment for Zta across every tested promoter compared to the OriLyt-Flank control region (no binding site) (**Figure 3-5 (A and B)**). Student t-test was used to calculate the P-values for the weak Zta binding at *BFRF2*, *BLLF2*, *BRRF1* and *BNLF2a* promoters compared to the OriLyt-Flank control region. Statistically significant Zta enrichment was observed in LCL#3 ( $P < 0.001$ ; marked with (\*\*\*)) and 293-BZLF1-ZKO ( $P < 0.01$  (marked with (\*\*)) (**Figure 3-5 A and B, respectively**).

Similar to our findings at immediate-early promoters, there was no significant variation in Zta association with the tested EBV early lytic promoters between the two EBV-positive cell lines used in this experiment.

**A****B**

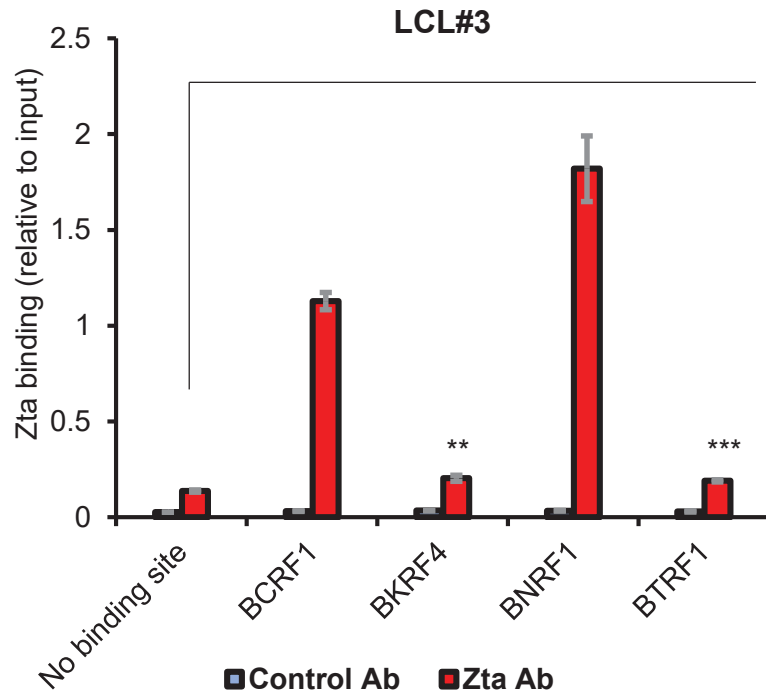
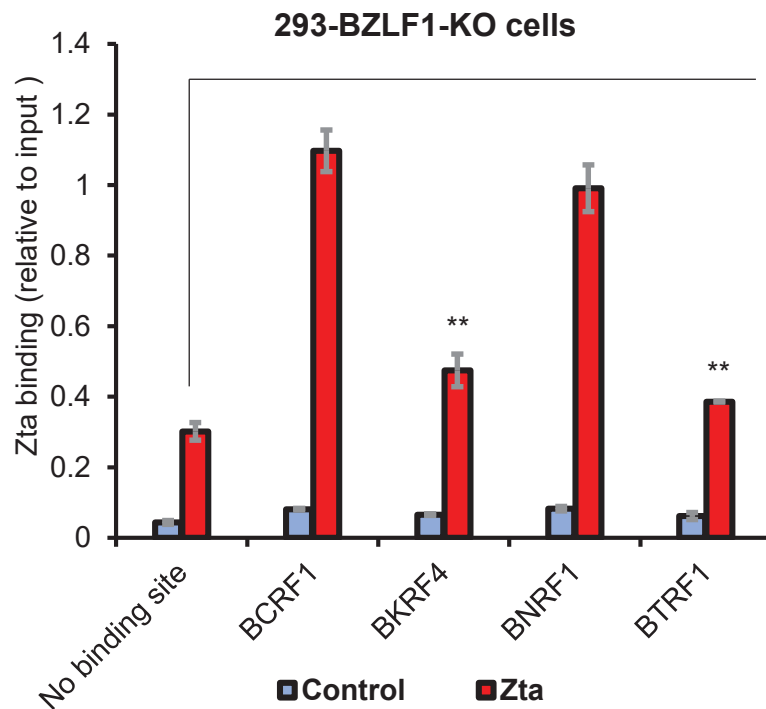
**Figure 3-5 Zta binding at the early lytic promoters.** ChIP-qPCR in different EBV cell lines: **(A)** LCL#3 cells and **(B)** 293-BZLF1-KO. Zta binding (red bars) is compared to a control sample (blue bars). The qPCR assays were performed using primers for no binding site (OriLyt-Flank) and various early lytic promoters. Values represent absolute levels of enrichment calculated using a standard curve of the input material. Error bars represent the mean of triplicate readings  $\pm$  SD for each sample. (\*\*\*) indicates  $p < 0.001$  while (\*\*) indicates  $p < 0.01$  compared to No binding site.

### 3.2.5 Zta binds similarly to promoter regions of various late lytic genes

The association between Zta and various promoters of EBV late lytic genes was also addressed in our ChIP-qPCR investigation. The late lytic promoters differ from the immediate-early and early lytic promoters in terms of their regulation, TATA box motif, and their temporal expression as described in **Chapter 1**. Zta association with some of these genes suggests a strong possibility that a different Zta mechanism is regulating these genes.

The qPCR data showed a significant enrichment for Zta across every tested promoter compared to the OriLyt-Flank control region (no binding site) (**Figure 3-6 (A and B)**). Student t-test was used to calculate the P-values for the weak Zta binding at *BKRF4*, and *BTRF1* promoters in comparison with the OriLyt-Flank control region. Significant Zta enrichment was observed in LCL#3 and 293-BZLF1-ZKO (**Figure 3-6 A and B, respectively**). P-values were ranging from  $< 0.001$  in (marked with (\*\*)), and  $< 0.01$  (marked with (\*\*)).

Our findings indicate that there was no significant variation in Zta association with the tested EBV late lytic promoters between the two EBV-positive cell lines used in this experiment.

**A****B**

**Figure 3-6 Zta binding at the late lytic promoters.** ChIP-qPCR in different EBV cell lines: **(A)** LCL#3 cells and **(B)** 293-BZLF1-KO. Zta binding (red bars) is compared to a control sample (blue bars). The qPCR assays were performed using primers for no binding site (OriLyt-Flank) and various late lytic promoters. Values represent absolute levels of enrichment calculated using a standard curve of the input material. Error bars represent the mean of triplicate readings  $\pm$  SD for each sample. (\*\*\*) indicates  $p < 0.001$  while (\*\*) indicates  $p < 0.01$  compared to the No binding site.



### 3.2.6 Zta binds to a novel site with unknown regulatory function

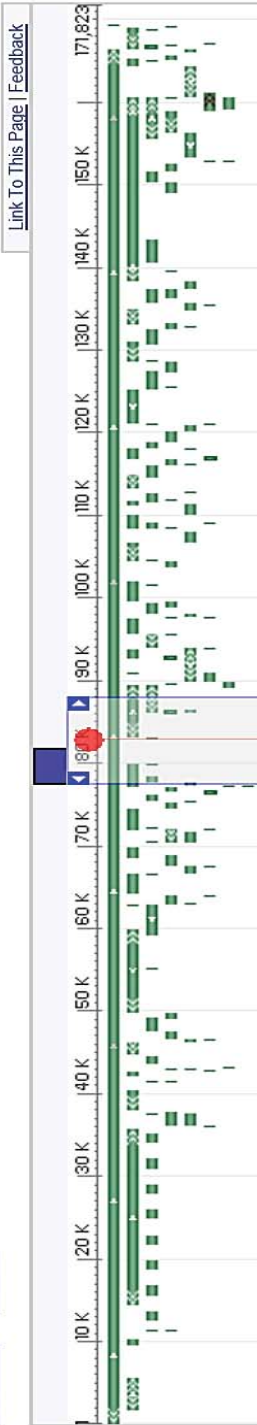
The ChIP-qPCR approach, as previously described, was used to evaluate the association between Zta and the wildtype EBV strain (NC\_007605) genomic region between 82462-82962 bp. Specific qPCR primers were designed, as listed in **Table 2-4** with their genomic coordinates. This region (referred to as ~82 kbp) was of interest because it is about 2 kbp away from the closest lytic open reading frame (*BLLF1* and *BLLF2*) (**Figure 3-7**). Furthermore, it was one of the primary regions we found to be different in the published Zta ChIP-Seq data as explained in **Figure 3-1**.

**Figure 3-7** represents a genome view of this ~82 kbp region taken from the publicly available NCBI genome browser. This view shows EBV coding genes in close proximity to ~82 kbp region, and the nearest TATA boxes (highlighted in red). A Promo web tool search for Zta binding sites (Farre et al., 2003) revealed two binding sites TGTCACA (82765-82771 bp) and TGTGTCA (82825-82831 bp). The location of these ZREs is shown in **Figure 3-7 (B)**.

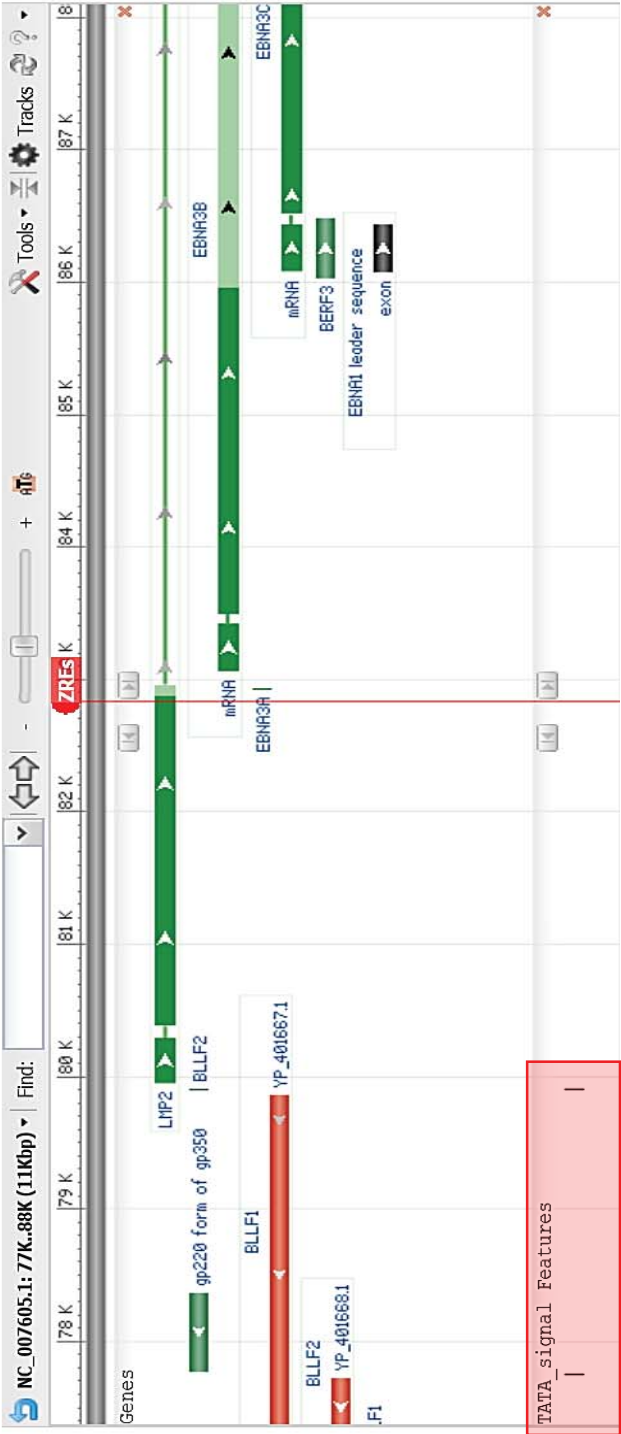
Zta binding was observed at this novel region in both cell lines tested in our experiment. Statistically significant enrichment for Zta was seen compared to the OriLyt-Flank control region. Student t-test was used to calculate the P-values which were < 0.001 in both cell lines (marked with (\*\*\*)) (**Figure 3-7 (C)**).

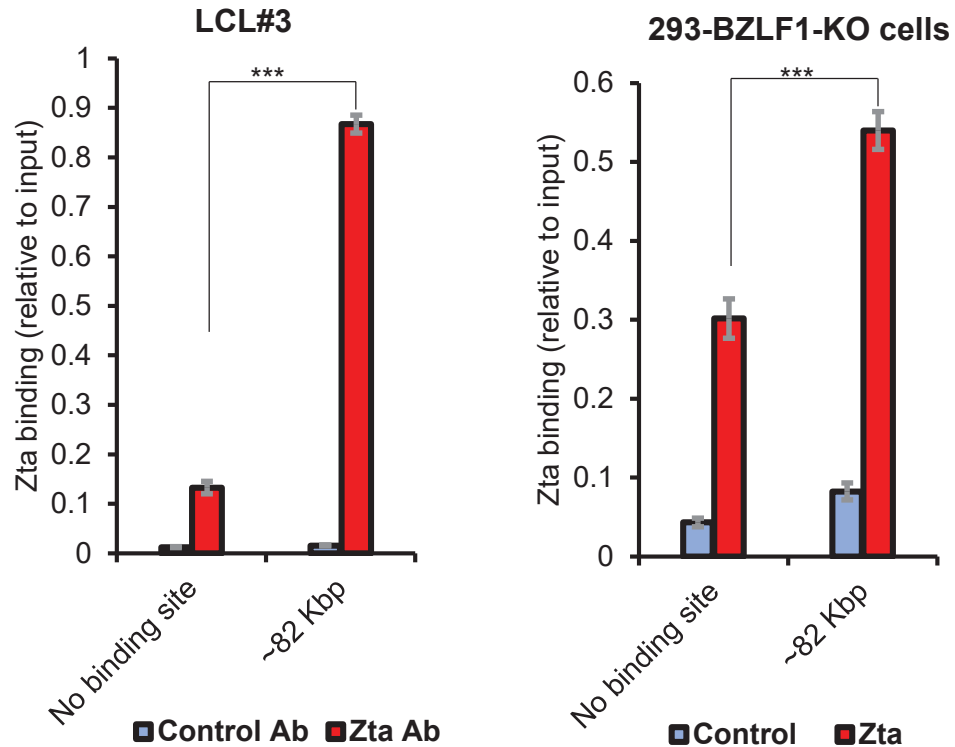
To investigate if there is any EBV sequence variation that might affect Zta binding to this region in Raji EBV strain, a BLAST analysis of the region between 82462-82962 bp (NC\_007605) was carried out in Akata EBV strain sequence (KC207813) and Raji EBV strain sequence (KF717093) (**Figure 3-8**). Although the predicted ZRE sites were identical in both strains, the region flanking one of the ZRE in Raji strain was different from the EBV wildtype strain. This suggests that EBV strain sequence variations might play a major role in Zta binding patterns.

**A** NCBI Reference Sequence: NC\_007605.1  
[GenBank](#) [FASTA](#)



**B**



**C**

**Figure 3-7 Zta binding at a region with unknown regulatory function (~82 kbp).** (A) NCBI genome browser overview of the wildtype EBV sequence (NC\_007605). The target region ~82 kbp is marked with a red flag; genes are shown as green bars. (B) A zoomed-in view of ~82 kbp showing the annotated ORFs in close proximity to this region. Red bars represent protein coding regions. This site maps to a region around 95 kb in the EBV B95-8 strain (V01555) as previously described in Figure 3-1. Nearby TATA box signals are shown as short black vertical lines (red box). The approximate location of the predicted ZREs (using PROMO web tool) in this target region is marked with a red flag (C) ChIP-qPCR in different EBV cell lines: LCL#3 cells (left) and 293-BZLF1-KO (right). Zta binding (red bars) is compared to a control sample (blue bars). The qPCR assays were performed using primers (Table 2-4) for no binding site (OriLyt-Flank) and ~ 82 kbp region. Values represent absolute levels of enrichment calculated using a standard curve of the input material. Error bars represent the mean of triplicate readings  $\pm$  SD for each sample. (\*\*\*) indicates  $p < 0.001$  compared to the No binding site.

**A**

Human herpesvirus 4 strain Akata, complete genome

Sequence ID: [KC207813.1](#) Length: 171323 Number of Matches: 3Range 1: 82130 to 82630 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
920 bits(498)	0.0	500/501(99%)	0/501(0%)	Plus/Plus

Query	241	TTGATTTGTCAATCCATGGCAGGCCCGCCCTCGGACCCCGAGTGGCCTGTTCAAGGGG	300
Sbjct	82370	TTGATTTGTCAATCCATGGCAGGCCCGCCCTCGGACCCCGAGTGGCCTGTTCAAGGGG	82429
Query	301	AAGGTGGCCAAAATGTCACAGGCCCTGAAACTAGAAGGGTGGTGGTGTGCTAGCTGTTGTTC	360
Sbjct	82430	AAGGTGGCCAAAATGTCACAGGCCCTGAAACTAGAAGGGTGGTGGTGTGCTAGCTGTTGTTC	82489
Query	361	ACATGTGTCAAGGATGACGAGTTTCCGGATCTACAAGATCCTCCAGATGAGGCCTAAGCAA	420
Sbjct	82490	ACATGTGTCAAGGATGACGAGTTTCCGGATCTACAAGATCCTCCAGATGAGGCCTAAGCAA	82549
Query	421	AGGTGTAGAAGTGTGTCCCCCTCCATTCCACCCACTGATAATACGCCCGACAATAAAGTT	480
Sbjct	82550	AGGTGTAGAAGTGTGTCCCCCTCCATTCCACCCACTGATAATACGCCCGACAATAAAGTT	82609

**B**

Human herpesvirus 4 strain Raji, complete genome

Sequence ID: [KF717093.1](#) Length: 166182 Number of Matches: 2Range 1: 82469 to 82969 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
854 bits(462)	0.0	488/501(97%)	0/501(0%)	Plus/Plus

Query	241	TTGATTTGTCAATCCATGGCAGGCCCGCCCTCGGACCCCGAGTGGCCTGTTCAAGGGG	300
Sbjct	82709	TTGATTTGTCAATCCATGGCAGGCCCGCCCTCGGACCCCGAGTGGCCTGTTCAAGGGG	82768
Query	301	AAGGTGGCCAAAATGTCACAGGCCCTGAAACTAGAAGGGTGGTGGTGTGCTAGCTGTTGTTC	360
Sbjct	82769	AGAGTGGCCAGAAATGTCACAGACCATGAACCTAGAAGGGTGGTGGTGTGCTAGCTATTGTTC	82828
Query	361	ACATGTGTCAAGGATGACGAGTTTCCGGATCTACAAGATCCTCCAGATGAGGCCTAAGCAA	420
Sbjct	82829	ACATGTGTCAAGGATGACGAGTTTCCGGATCTACAAGATCCTCCAGATGAGGCCTAAGCAA	82888
Query	421	AGGTGTAGAAGTGTGTCCCCCTCCATTCCACCCACTGATAATACGCCCGACAATAAAGTT	480
Sbjct	82889	AGGTGTAGAAGTGTGTCCCCCTCCATTCCACCCACTGATAATACGCCCGACAATAAAGTT	82948

**Figure 3-8 Sequence variations between EBV strains at the ~82 kbp region.** NCBI nucleotide BLASTs between the region 82462-82962 of the wildtype EBV sequence (NC\_007605) and the complete sequence of EBV Akata (KC207813) and Raji (KF717093) strains were carried out. **(A and B)** show a small region around the predicted ZREs of the BLAST results in ~82 kbp region. Akata EBV shows 99% similarity at this region **(A)** while Raji EBV strain has only 97% sequence similarity at this region **(B)**. The predicted ZREs in ~82 kbp region are shown in red boxes.

### 3.3 Discussion

Zta plays a critical role in triggering the switch between latency and lytic phase by initiating viral DNA replication and regulating various lytic and host genes as previously described. There has been much interest in establishing the binding sites for Zta across the viral and cellular genome using genome-wide approaches such as ChIP-Seq analysis.

Although this highly sensitive technique has revealed a wealth of information regarding Zta binding patterns, the overall view of Zta binding is still lacking. This is mainly due to the fact that there are different EBV systems with different EBV strains. Not to mention that to date there is no global Zta binding study published in an EBV epithelial system. Moreover, our observation of differences between the published Zta ChIP-Seq data sets in Burkitt's lymphoma B cells across the viral genome (Bergbauer et al., 2010; Ramasubramanyan et al., 2012a) has raised many questions about Zta binding in different cells, as described in **Section 3.2.1**.

In this chapter, I carried out ChIP-qPCR analysis to compare Zta binding across the viral genome in two different EBV cell lines: the spontaneously lytic LCL#3 (Sinclair et al., 1994a), and the latent 293-BZLF1-KO cells (Feederle et al., 2000) chosen as a non-B cell EBV system.

Unlike ChIP-Seq, the ChIP-qPCR approach has a low resolution and is limited to individual target sites. However, it is considered as a very efficient, rapid and cost-effective way to investigate transcription factors (TFs) binding as well as histone modifications *in vivo* (Gade and Kalvakolanu, 2012).

Using this approach, I aimed to address the variations observed in published ChIP-Seq data described earlier. In addition, I considered validating Zta binding at novel EBV promoter sites identified previously by members of our lab (Ramasubramanyan et al., 2012b).

The two ChIP-Seq data sets previously described, by Bergbauer et al and Ramasubramanyan et al were carefully aligned to identify any differences in Zta

binding. Both data sets were largely in agreement; especially at major lytic promoters (*BZLF1*, *BRLF1*, *BALF5*, and *BMRF1*). However, Zta binding was clearly different at a few sites. Two examples of these differences are provided in **Figure 3-1**. One of these sites was the promoter of the late lytic gene *BNRF1*, which is one of the major tegument proteins (Cameron et al., 1987).

These differences in Zta binding ChIP-Seq published data could be explained by many factors including technical and biological variations. Technical variations in ChIP-Seq can be due to differences in experimental conditions such as antibodies, chromatin fragmentation, quality of the input DNA, and peak calling methods. Such variations may introduce bias which makes it difficult to compare data from different labs. Nevertheless, ChIP-Seq remains a very powerful and sensitive technique that produces consistent results (Haring et al., 2007; Kidder et al., 2011).

On the other hand, these differences can occur due to biological variations, specifically DNA sequence variations between EBV strains, host cell specific variations and viral chromatin structure. For example, Raji BL EBV strain harbors a deletion affecting *BALF2*, which results in defective EBV replication (Decaussin et al., 1995; Hatfull et al., 1988). It is worth noting that Zta was shown to bind to the *BZLF1* promoter in B95-8 but not in Raji cells, suggesting that cell or EBV genome variations can influence *BZLF1* promoter (Zp) regulation (Bergbauer et al., 2010). Also, *BZLF1* promoter is not activated through the phorbol ester TPA or histone deacetylase (HDAC) inhibitors in Raji cells (Murata et al., 2012). Although Raji EBV strain has the sequence of the prototype Zp, many sequence variants of Zp were described in other EBV strains (Feederle et al., 2015).

Also, cell specific variation in chromatin regulation can influence the chromatin structure of EBV. The Akata EBV genome has decreased histone acetylation while Raji EBV genome is associated with high H3k27me3 and H4K20me3, reviewed in (Lieberman, 2015).

B cells have been predominantly used to investigate the epigenome of EBV. On the contrary, only limited promoter regions were investigated in EBV epithelial systems such as in 293 EBV-BAC cells (Murata et al., 2012).

In addressing these variations, another factor that is worth considering was the replacement of Zta transactivation domain by a GFP protein in the ChIP-Seq experiments carried out by Bergbaur et al. This GFP-Zta construct included amino acids between 149-245 (Kalla et al., 2010) which were shown to have a very similar affinity to DNA binding compared to full-length Zta *in vitro* (Flemington et al., 1992). However, DNA binding differences between the full-length Zta and GFP chimeric bZIP Zta cannot be ruled out *in vivo*. Zta transactivation domain may recruit other partner proteins that are required to stabilise or change the chromatin state at certain sites.

In highly lytic EBV systems, Zta can bind to targets across the newly replicated EBV genome copies, which are not chromatinised. This could also introduce bias in the data. Despite this, similar findings for Zta binding were reported under replicated and non-replicated conditions in Akata cells (Ramasubramanyan et al., 2012b).

In our ChIP-qPCR experiments, to cover a wide range of targets, various lytic promoters were investigated using a similar protocol and Zta antibody used in the experiments carried out by Ramasubramanayn et al. These target regions include promoters for genes classified as immediate-early, early, and late lytic genes. Investigating these different promoters was important because different mechanisms for each class is postulated, as described in **Chapter 1**.

To date most of these target promoters have not been investigated for their response to Zta transcriptional regulation with the exception of *BZLF1* (Flemington and Speck, 1990a; Lieberman and Berk, 1990), *BRLF1* (Sinclair et al., 1991), *BRRF1* (Segouffin-Cariou et al., 2000) and the late lytic cycle gene *BCRF1*, a homolog to cellular *IL10* (Mahot et al., 2003b). In addition, in the last



two chapters, our analysis will show that *BNLF2a* is a new target regulated by Zta.

Furthermore, the ChIP-qPCR was performed in two different EBV systems as described earlier. LCL#3 cells are predominantly in latency type III but contain a small sub-population of cells that are continuously in lytic form (Bernig et al., 2014; Sinclair et al., 1994a). The 293-BZLF1-KO cells used in this experiment carry a *BZLF1* mutant EBV. Thus, to induce lytic in these cells, transient transfection with Zta was required. Cells were transfected with His-tagged Zta and harvested 48 hours after transfection. In this system, DNA replication was previously shown to start around 96 hours post Zta transfection (Bailey et al., 2009a). In theory, the number of EBV infected cells that go into lytic cycle in this system will be dependent on Zta transfection efficiency. Although, Zta expression was checked by western blots in 293-BZLF1-KO cells, none of the two cell lines were analysed for the percentage of cells that contained Zta expression, and hence supported EBV lytic induction.

As was shown in **Figure 3-3 - Figure 3-7**, Zta binding patterns were consistent between both cell lines tested, and in agreement with the data reported in Akata BL cells (Ramasubramanian et al., 2012b). Across the various analysed EBV regions, Zta enrichment was varied. The strongest enrichment was observed within the OriLyt site, which has 7 ZREs (Schepers et al., 1996). It is not clear if this enrichment is related to the number of binding sites, as stronger Zta binding was also observed in *BNRF1* promoter which contains two ZREs, whereas weaker Zta enrichment in other regions contained as many as five ZREs. Thus, it is very likely that these regions contain low-affinity sites or have open chromatin organization that influences Zta binding (Landt et al., 2012).

Interestingly, Zta binding was validated at a region, which contains ZREs, of unknown regulatory function (~82 kbp) as described in **Figure 3-7**. Indeed, several other regions for Zta binding at uncharacterised genome sites have been reported (Ramasubramanian et al., 2012a). These regions (including ~82 kbp) could be possibly a proximal enhancer element for a gene nearby or play a role in Zta genome replication. Similar regions with unknown functions have



also been shown to be highly co-occupied with several of TFs during latency. For example, ATF3 and USF TFs were reported to co-localize at the region around 80,655 bp. This site is located within the first EBNA3A transcript internal repeat (Arvey et al., 2013). A Zta binding peak is also present at this site in the examined ChIP-Seq data by Ramasubramanyan et al and Bergbauer et al. Unlike the other regions, Zta binding peak at ~82 kbp was not seen in the Raji cell line ChIP-Seq data by Bergbauer et al which suggests this site is very likely bound by Zta in a cell line/ EBV strain -specific manner.

The sequence of this region in Raji strain was only 97% similar to wildtype EBV strain. The region flanking one of the predicted ZREs in the ~82 kbp was affected by these variations (**Figure 3-8**). This could affect Zta binding or other transcription factors required for Zta binding at this site.

In conclusion, Zta binding to the EBV genome is influenced by many factors including sequence variations, viral epigenetic structure, and cell-specific factors. Although the differences between Zta binding in the published data will require more analysis, I was able to extensively assess Zta binding across the viral genome in two completely different EBV systems that were largely unused and uncharacterized for Zta binding analysis during EBV lytic cycle.

## **Chapter 4 ZREs as long-distance enhancer elements**

### **4.1 Introduction**

In eukaryotic cells, gene expression is dependent on the spatiotemporal binding of regulatory proteins (e.g. transcription factors (TFs)) to promoter and enhancer/silencer elements. This takes place within a very complex dynamic 3D chromatin organization that involves the recruitments of transcription factors, engagement of chromatin remodelers and modifiers, and interactions between different regulatory elements. These complex processes have one main goal of recruiting the general transcription machinery (including Pol II) to the core promoter of a specific gene. This in turn leads to the final outcome which is the initiation of mRNA transcription and eventually protein production.

Understanding this 3D organization of the genome has attracted much interest in the last few years. Recent advances in genome-wide technologies facilitated the deciphering of the complex transcription regulation mechanisms in various cells as well as organisms (Chen et al., 2013; Gerstein et al., 2012; Homouz and Kudlicki, 2013; Sanyal et al., 2012; Shen et al., 2012). Furthermore, these studies have integrated data from various resources and techniques such as ChIP-Seq, chromosome conformation capture methods, and RNA-Seq.

To be more specific, we gained more insight into how promoters and long-range enhancer elements interact in such a complex chromatin environment. The first observation of enhancer activity was observed when a 72 bp repeated sequence taken from the SV40 virus was inserted into a recombinant DNA plasmid expressing rabbit haemoglobin beta 1 gene. This SV40 DNA resulted in an approximately 200-fold change in the transcript levels of that gene (Banerji et al., 1981).

Enhancers are regulatory elements that typically span a few hundred base pairs, with clusters of different transcription factors (TFs) binding sites. Their main role is to enhance transcription of a target gene or genes in spite of their orientation and distance relative to their targets (Spitz and Furlong, 2012). An active enhancer is characterised by open chromatin structure. This is indicated

by the low nucleosome occupancy in the enhancer region. Also, nearby nucleosomes are commonly associated with the following histone modifications: enrichment of monomethylation of histone H3 lysine 4 (H3K4me1) and the acetylation or methylation of histone H3 lysine 27 (H3K27). Moreover, CREB-binding protein (CBP) and its related protein p300, which have histone acetyltransferase activity and known to interact with many other TFs and histone modifiers, have been shown to be enriched at enhancer elements. On the contrary, inactive enhancer elements are associated with histone modification H3K27me3, a polycomb protein-associated repressive mark, and can be bound by suppressor TF binding, (reviewed in (Ong and Corces, 2011; Shlyueva et al., 2014).

Enhancers play a key role in regulating cell-specific expression which is important in determining cell differentiation and fate (Heintzman et al., 2009; Visel et al., 2007). Also, mutations in enhancer elements have been associated with the development of many cancers. A well-known example is Burkitt's lymphoma, a tumour that is associated with EBV. A translocation that brings the *c-Myc* oncogene under the control of the enhancer region of the immunoglobulin heavy chain in B cells, leads to strong overexpression of c-Myc (Dalla-Favera et al., 1982) and eventually the development of Burkitt's lymphoma, as described in **Section 1.3.6**. The role of enhancers in various cancers is reviewed in (Herz et al., 2014; Sur and Taipale, 2016).

Recently, the term 'super enhancers' have been used to describe atypical enhancers. These enhancers are usually associated with a very strong enrichment of the master transcription factors in ESCs: Oct-4 Sox2, Nanog, KLF4 and Esrrb and were initially identified in embryonic stem cells (Whyte et al., 2013). In contrast to typical enhancers, a distinguishing feature of these super enhancers is the fact that they span a long region that is enriched for one transcription coactivator, Mediator (Med1). This association with the mediator coactivator complex facilitate the interaction with the general transcription factors and Pol II (Whyte et al., 2013). Further examination of the occupancy of these super enhancer sites in different cell types, has associated these elements with a wide range of other transcription factors in a cell type specific

manner (Hnisz et al., 2013). The intriguing subject of super enhancers is recently reviewed in (Pott and Lieb, 2015).

Enhancers are not only bound by cellular factors, but many viral proteins have been implicated in gene regulation through enhancers. Recent studies described the involvement of EBV proteins in enhancers and super enhancer region elements. In particular, EBNA2 has been associated with the regulation of various oncogenes such as *c-Myc* and *BCL2* in B cells (Gunnell et al., 2016; McClellan et al., 2013; Zhou et al., 2015).

A recent Zta genome-wide study conducted in our lab using ChIP-Seq and RNA-Seq in EBV Burkitt's lymphoma Akata cells revealed that approximately 75% of the Zta binding peaks were at distal elements (>4 kb) from the transcription start sites (TSS) of cellular genes across the genome, while only 15% were within 2 kb of TSS (Ramasubramanyan et al., 2015). Moreover, addressing the functional relevance of Zta binding, the ChIP-Seq data were integrated with RNA-Seq data which indicated that around 278 genes associated with Zta binding showed change in expression upon Zta induction in these cells. Zta binding peaks were observed at variable locations across these genes. For example, in some genes such as *RASA3* and *FOSB*, Zta binding peaks were found at various long distance locations from the genes including 5', and 3' and intragenic sites (Ramasubramanyan et al., 2015).

Based on these data, we hypothesised that Zta regulation is possibly mediated through long-range enhancer elements. In order to test this hypothesis, we have designed a synthetic DNA fragment with 20 CpG-free ZREs (class I), placed around 2 kb upstream a heterologous minimal promoter in a luciferase reporter construct.

## 4.2 Results

### 4.2.1 Luciferase reporter assay to investigate ZREs as potential enhancer elements

Given the recent finding of Zta binding at distal elements within the cellular genome (Ramasubramanyan et al., 2015) as discussed in **Section 4.1**, we took a simple approach to evaluate ZREs as potential long-range enhancer elements, outlined in **Figure 4-1**.

A minimal promoter was subcloned into the pGL3-Control luciferase reporter plasmid between KpnI and HindIII RE sites, upstream of the luciferase gene. In the same construct, a random 20X ZREs synthetic DNA fragment, with similar spacing between each ZRE, was subcloned into the second multiple cloning site of pGL3-Control, between BamHI and Sall. This second multiple cloning site in the pGL3 vector is around 2.2 kb away from the promoter and designed to test potential enhancer elements. As shown in **Figure 4-1**, a schematic diagram of the designed construct is given in **(A)** (dark grey plasmid). The minimal promoter is shown in orange while the 20XZRE construct is shown in blue.

Two constructs (MinP and MinC) were generated using two different promoters along with the same 20XZRE insert. This is further illustrated in **Figure 4-2 (A and B)**.

The luciferase constructs were co-transfected with either a control vector (pcDNA3) or His-Zta expression vector in EBV-negative Burkitt's lymphoma cell line (DG75). This allowed the comparison of luciferase activity in the presence and absence of Zta as demonstrated in **Figure 4-1 (B)**.

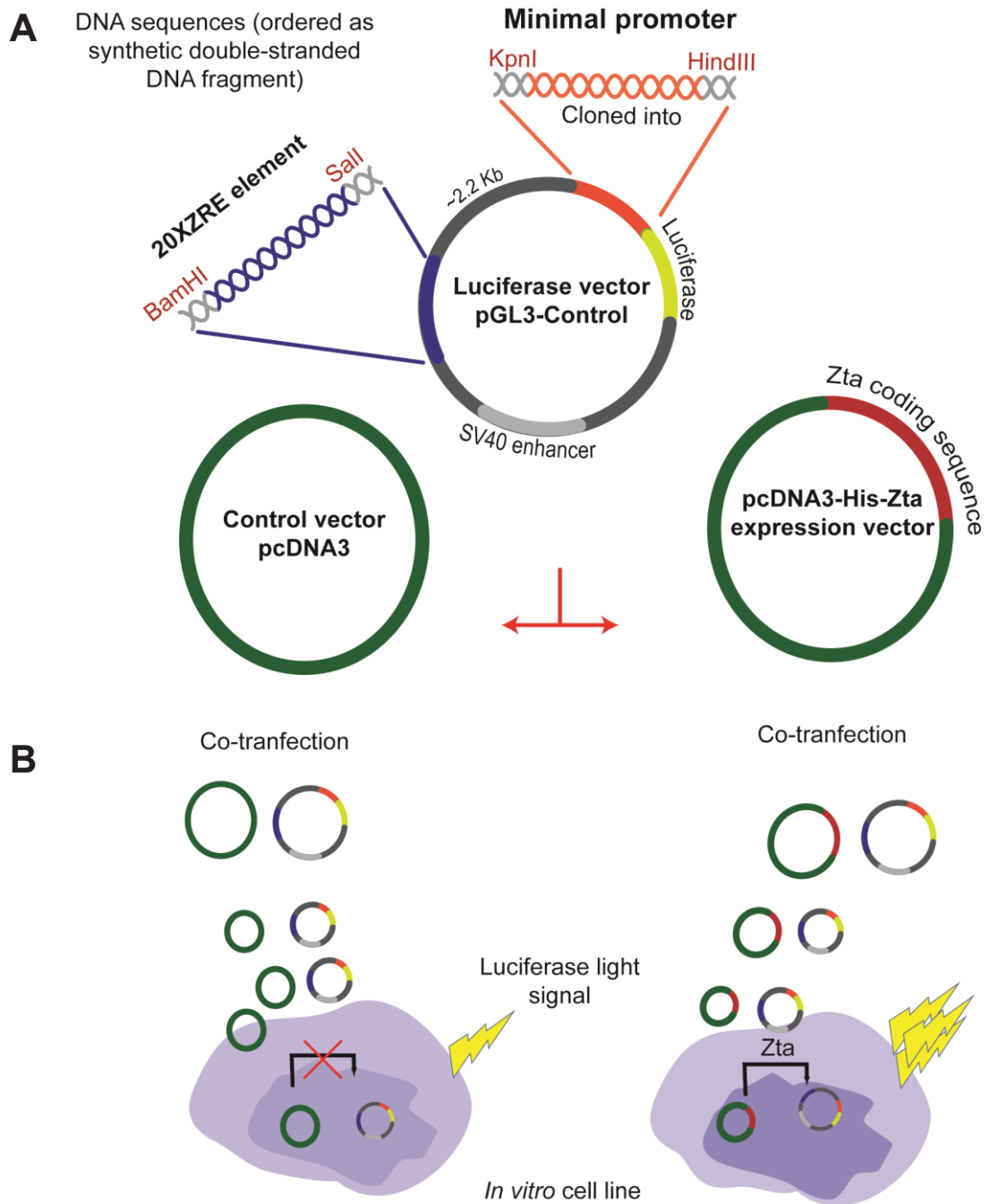
**Figure 4-2** represents schematic diagrams as well as the exact sequences of the minimal promoters and 20XZRE inserts for both luciferase constructs (MinP and MinC). The sequence of the minimal promoter (**Figure 4-2 (A)**) in MinP construct was adapted from the minimal promoter in Promega luciferase construct (pGL4). This 32 bp promoter was ordered as a synthetic double-

stranded DNA fragment after the addition of compatible restriction enzyme (RE) sites at each end to facilitate the subcloning into pGL3-Control plasmid.

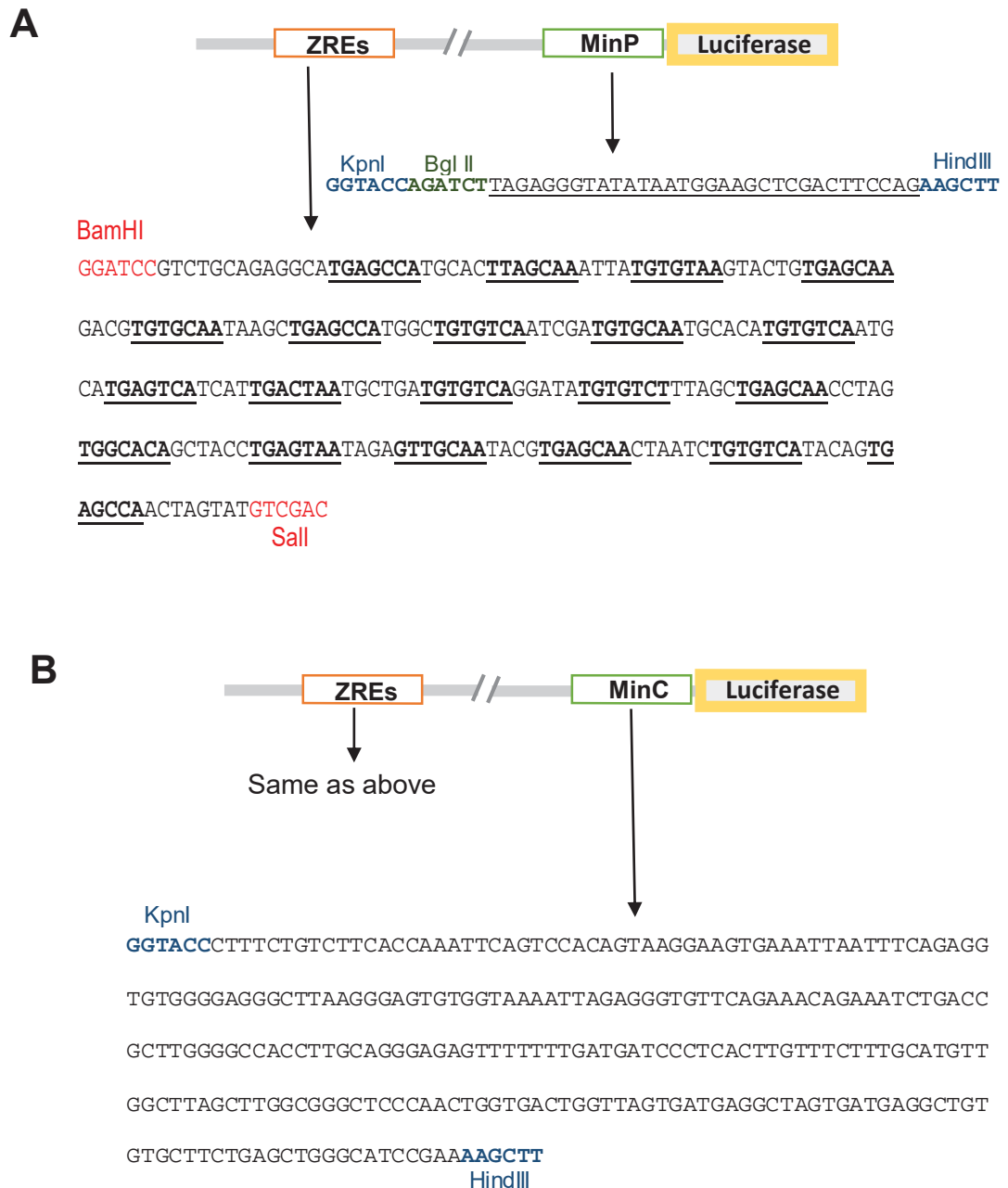
Also, the distal 20XZRE element sequence is shown in **Figure 4-2 (A)** where each ZRE motif is underlined. The 20 ZREs used in this synthetic insert were class I ZREs, which do not contain a CpG methylation motif, and were previously described in various Zta target promoters (Flower et al., 2011).

**Figure 4-2 (B)** is a schematic diagram of MinC construct and the MinC promoter sequence. This construct differs from the MinP construct in the promoter sequence. This second construct was generated to ask if the same effect of Zta in MinP was reproducible on a different promoter. The MinC promoter is part of the cellular gene *CIITA* promoter (*CIITA* -214+53) and was previously cloned into pGL3 luciferase vector by Dr Nicolae Balan in our lab (Balan et al., 2016).

The results of luciferase reporter assay using MinP and MinC, shown in **Figure 4-3** and **Figure 4-4**, have been recently published (Ramasubramanyan et al., 2015).



**Figure 4-1 Investigating a long-range ZRE element effect on promoter activation.** A schematic representation of luciferase reporter assay experiment carried out to investigate Zta activation through a long-range enhancer elements. **(A)** A synthetic DNA fragment containing 20 ZREs (blue) was cloned between BamHI and SalI restriction sites in pGL3-Control vector (Promega). The ZRE fragment is around 2 kb away from a minimal promoter (orange). Two different minimal promoter constructs (MinP or MinC) were generated as will be discussed in Figure 4-2. Promoter sequences were cloned between KpnI and HindIII upstream the luciferase expression sequence. **(B)** The luciferase vector was co-transfected with either a control vector (pcDNA3) (green plasmid) or His-Zta expression vector (green and red plasmid) in EBV negative Burkitt's lymphoma cell line (DG75). Luciferase signal was compared in the absence and presence of Zta.



**Figure 4-2 DNA sequences of the 20XZRE element and heterologous minimal promoters. (A)** A schematic representation of MinP luciferase construct. The sequence and cloning restriction enzyme sites (REs) of the heterologous minimal promoter are shown. The minimal promoter is a 32 bp promoter containing a TATA-box motif taken from Promega pGL4 luciferase vector. The 2.2 kb distal 20XZRE element sequence is shown along with the cloning restriction enzyme sites (red) where each ZRE is underlined. **(B)** A schematic representation of MinC luciferase construct which is similar to (A) but has a different promoter taken from the host *CIITA* gene (referred to as MinC). The sequence and cloning REs sites of the heterologous minimal promoter (*CIITA* -214+53) as well as the 20XZRE element are shown below the diagram. *CIITA* -214+53 is a known target for Zta shown previously in our lab (Balan et al., 2016). The 20X ZRE element was cloned into MinC by Ms Renu Gurung (a final year undergraduate student taking her project under my supervision in our lab).



#### 4.2.2 Zta activates a heterologous minimal promoter through distal ZREs

As described in **Section 4.2.1**, a luciferase construct (MinP) was generated to test the potential of ZREs to act as potential long-range enhancer elements. The luciferase activity of this construct was compared in the presence and absence of Zta expression in EBV-negative DG75 cells.

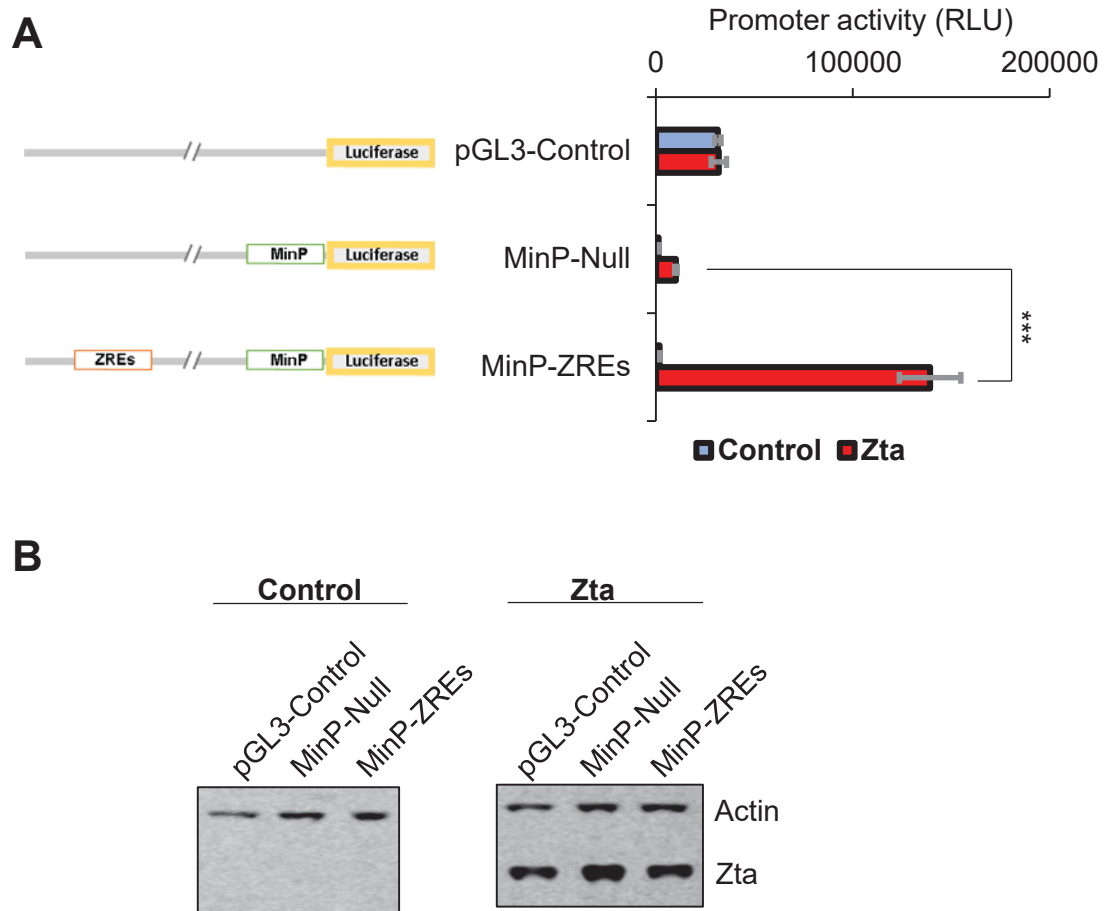
**Figure 4-3 (A)** represents a luciferase reporter experiment showing a comparison of the promoter activity, in normalised luciferase relative light units, for three constructs. These three constructs are the pGL3-Control luciferase plasmid (empty vector), MinP-Null (construct with heterologous minimal promoter, but no distal 20XZRE element), and MinP-ZREs (construct with MinP minimal promoter and distant 20XZRE element).

Comparison of the luciferase activity for each construct in the presence and absence of Zta showed approximately 73-fold change for MinP-ZREs compared to the control sample (in the absence of Zta). On the contrary, a 7 fold-change was seen for the MinP-Null construct. It is worth noting that the empty luciferase vector (pGL3-Control), which contains the SV enhancer element, showed no significant change in response to Zta.

A statistically significant 14-fold increase in MinP-ZREs luciferase activity compared to MinP-Null construct in the presence of Zta was seen (P-value <0.001).

**Figure 4-3 (B)** represents a western blot analysis for Actin (loading control) and Zta for each sample. Comparable levels of Zta expression can be seen for all samples while the absence of Zta is confirmed in the control samples. For each sample, the intensity of the actin band was used to normalise the luciferase relative light units value for that sample as previously described in **Chapter 2**.

The data presented in **Figure 4-3** strongly supports the emerging theory of a role for Zta and ZREs in long-range enhancer elements.



**Figure 4-3 A long-range ZREs element drives Zta activation of a heterologous minimal promoter. (A)** A comparison of the promoter activity for three constructs: pGL3-Control luciferase plasmid (empty vector), MinP-Null (construct with heterologous minimal promoter, MinP, but no distant 20XZRE element), and MinP-ZREs (construct with MinP minimal promoter and distant 20XZRE element). Constructs features and sequences are previously shown in Figure 4-2. Activities are compared in the presence (red bars) and absence (blue bars) of Zta expression in EBV-negative DG75 cells. Values are presented as the mean of relative light units (RLU) for triplicate readings  $\pm$  SD normalised to actin levels. **(B)** Western blot analysis of the samples showing similar levels of Zta transfection. (\*\*\*) indicates  $p < 0.001$  compared to MinP-Null in the presence of Zta.

### 4.2.3 Zta activates a known target promoter through distal ZREs

To confirm our previous findings, we generated a different luciferase construct (MinC) using a cellular Zta target promoter (*C/ITA*), as previously described in **Section 2.2.1**.

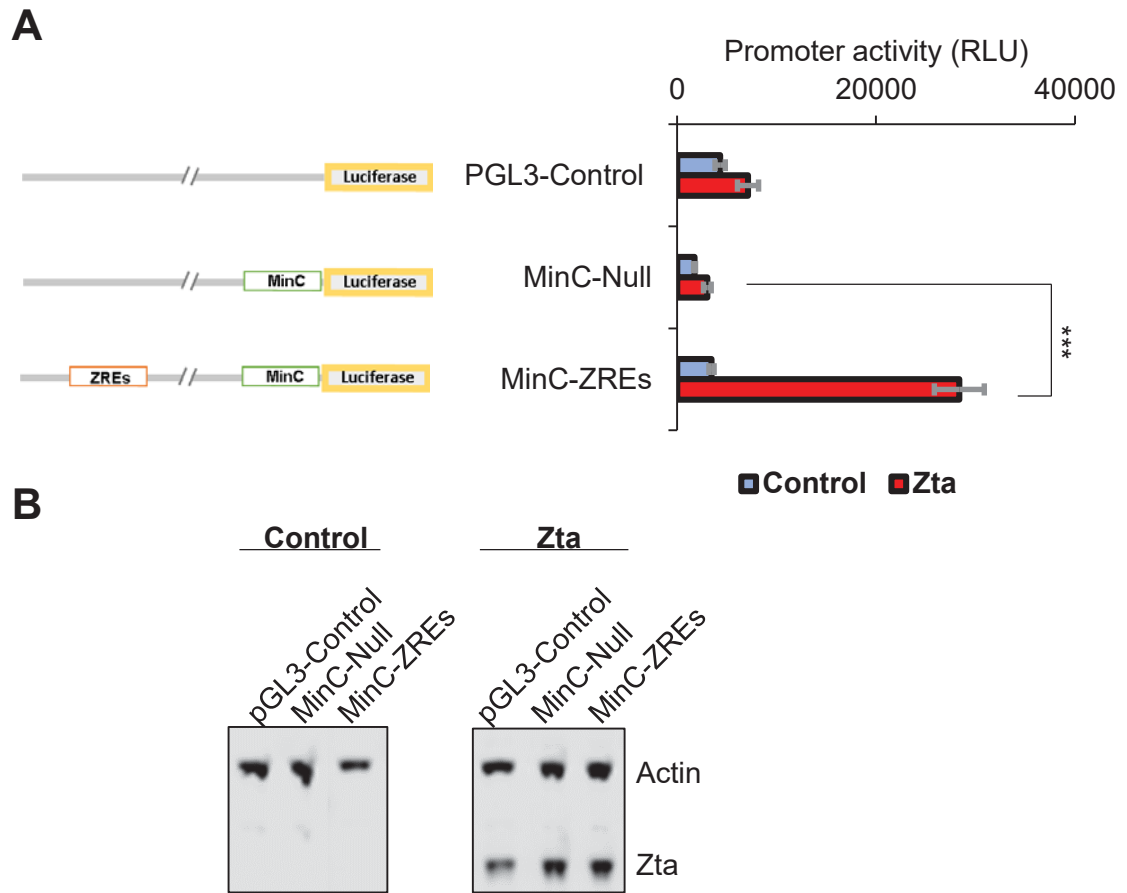
Largely similar to the previous experiment in **Figure 4-3**, the luciferase activity of the MinC construct was investigated in the presence and absence of Zta expression in EBV-negative DG75 cells (**Figure 4-4**).

**Figure 4-4** shows a comparison of promoter activity, in normalised luciferase relative light units, for three constructs: pGL3-control luciferase plasmid (empty vector), MinC-Null (construct with *C/ITA* minimal promoter, but no distant 20XZRE element), and MinC-ZREs (construct with *C/ITA* minimal promoter and distant 20XZRE element).

A statistically significant 9-fold increase in luciferase activity was seen in MinC-ZREs compared to MinC-Null construct in the presence of Zta (P-value <0.001).

**Figure 4-4 (B)** represents a western blot analysis for actin (loading control) and Zta for each sample. Comparable levels of Zta expression can be seen for all samples while the absence of Zta is confirmed in the control samples. For each sample, the intensity of the actin band was used to normalise the luciferase relative light units value for that sample as previously described in **Chapter 2**.

The data presented in **Figure 4-4** shows consistency in the 20XZRE enhancer element response to Zta expression under different promoters.



**Figure 4-4 A long-range ZRE element drives Zta activation of a known Zta target promoter. (A)** A comparison of the promoter activity for three constructs: pGL3-Control luciferase plasmid (empty vector), MinC-Null (construct with MinC minimal promoter but no distant 20XZRE element), and MinC-ZREs (a construct with heterologous minimal promoter and distant 20XZRE element). Constructs features and sequences are shown previously in Figure 4-2. Activities are compared in the presence (red bars) and absence (blue bars) of Zta expression in EBV-negative DG75 cells. Values are presented as the mean of relative light units (RLU) for triplicate readings  $\pm$  SD normalised to actin levels. **(B)** Western blot analysis of the samples showing similar levels of Zta transfection. (\*\*\*) indicates  $p < 0.001$  compared to MinP-Null in the presence of Zta.

### 4.3 Discussion

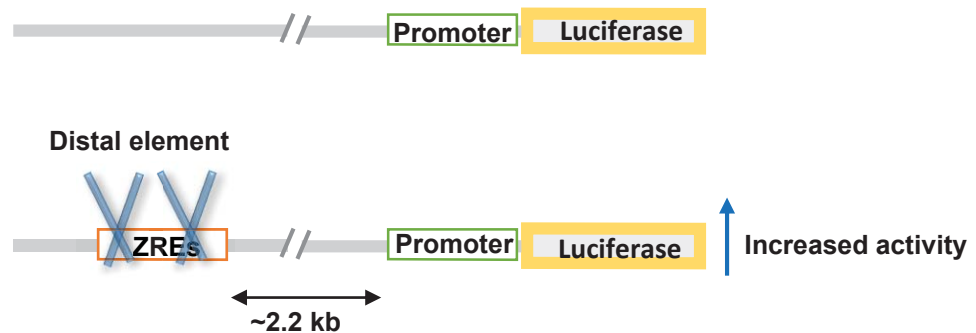
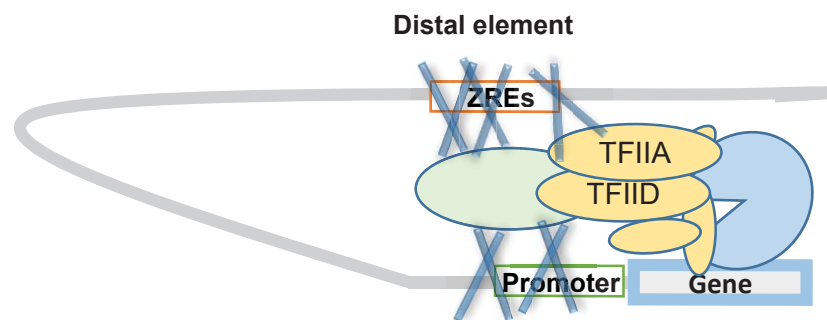
Unlike in the compact EBV genome, previous ChIP-Seq analysis carried out across the human genome in our lab showed that Zta binds at distal elements (> 4 kb), and these binding sites formed the majority of the total binding sites observed. Interestingly, this suggests a new possible regulatory role for Zta through long-range enhancer elements. Thus, to address this question, I used a simple and rapid luciferase reporter approach to investigate the ability of Zta to enhance the activation of a heterologous promoter *in vitro*. A luciferase construct, pGL3-Control (Promega, USA), designed to test the role of enhancer elements, was used in this approach. This vector has two cloning site regions allowing the subcloning of elements immediately upstream of the luciferase gene as well as another element further away.

A synthetic element containing tandem repeats of 20 different ZREs separated by short spacing was subcloned into pGL3-Control vector at a distance from a minimal promoter. Synthetic DNA fragments with random multi-ZREs have been used previously to investigate Zta regulatory mechanisms (Lieberman and Berk, 1991). In addition, a similar approach was used to analyse a novel methylated AP-1 site (mAP-1) bound by c-Jun/c-Fos heterodimer. A synthetic construct containing random five individual sites mAP-1 was cloned upstream of a minimal promoter in a luciferase reporter construct (Gustems et al., 2014).

The number of tandem repeat sites selected was considered to be the best reflection of enhancer elements. Commonly, enhancers are bound by multiple transcription factors where each transcription factor has multiple binding sites (Berman et al., 2002). The number of these multiple sites is variable; for example, in *Drosophila*, an enhancer element (eve stripe 2) was found to contain binding sites for at least four TFs, one of these TFs (krupple) has six multiple binding sites (Ludwig et al., 1998). Also, in SV40 enhancer, dissecting this element into shorter segments has resulted in the loss of their enhancer ability. However, when multiple copies of one shorter segment were used together, enhancer activity was restored (Schirm et al., 1987). Similarly, super enhancers do not only contain clusters for TFs binding sites but are arranged as multiple elements within a long enhancer region (Pott and Lieb, 2015).

Our *in vitro* experiments showed that the 20XZRE fragment mediated enhancer effects on both minimal promoters (MinP and MinC) as shown in **Figure 4-3** and **Figure 4-4**, respectively. Comparable levels of enhancer activity were observed for enhancer constructs ranging between 9- to 14-fold increases compared to the Null constructs in the presence of Zta. The fold change in activity levels were higher in the MinP than MinC construct when compared to the control sample lacking Zta expression (pcDNA3 control). This is attributed to the very low baseline activity of the MinP construct. The experiments were carried out in EBV-negative Burkitt's lymphoma (DG75). These cells were selected to eliminate any role for other viral TFs in this enhancer activity and to be in agreement with Zta ChIP-Seq data (described earlier) which was performed in Akata Burkitt's lymphoma cells.

Our luciferase reporter findings (summarised in **Figure 4-5 (A)**) provided functional evidence for Zta distance enhancement activity *in vitro* for the first time. However, a definitive evidence of Zta activation through long-range enhancer elements will require *in vivo* genome editing of enhancer elements as well as evidence of enhancer interaction with Zta regulated promoters. As shown in **Figure 4-5 (B)**, we propose a possible DNA looping model, reviewed in (Krivega and Dean, 2012), to explain how Zta achieves its regulatory role through enhancers. This could be facilitated by Zta direct binding to an enhancer element and a target promoter, which subsequently initiate a DNA loop and bring the two regions in close proximity. Zta binding partner such as the acetyl transferase coactivators CBP/p300 (Adamson and Kenney, 1999; Deng, 2003; Zerby et al., 1999) are very likely to be involved in this process. Also, as Zta is known to interact directly with TFIID and TFIIA (Lieberman, 1994; Lieberman et al., 1997), this could suggest a direct interaction between Zta and the general transcription factors at the core promoter in the absence of Zta binding sites within the promoter element as illustrated in our minimal promoters. Currently, the DNA loop model is being investigated in our lab using chromosome conformation capture (3C) method by Ijuel B. Naranjo Perez-Fernandez. This will provide new insights into Zta gene expression regulatory mechanisms.

**A****B**

**Figure 4-5 Increasing evidence suggesting a role for Zta at enhancer elements.** Zta binding patterns at vast distal regions across the human genome (Ramasubramanyan et al., 2015) as well as the increased luciferase promoter activity seen when a distal ZRE element was introduced to the pGL3-Control construct suggest a role for Zta in transcription regulation through enhancer elements. **(A)** A schematic diagram depicting a model summarising the long-range ZRE element experiment findings. **(B)** A possible DNA looping model of how Zta might facilitate the communication between the enhancer and promoter regions. Zta is likely to promote the interaction of long-range enhancer elements with core promoters either through the direct interaction with the basic transcriptional factors or through coactivators (especially in the presence of Zta response elements within the proximal promoters). Zta is shown as (blue X), co-activators (green oval), RNA Pol II and basal transcription machinery proteins are shown as blue circle, and yellow ovals, respectively.

## Chapter 5 Zta activates viral immune evasion gene *BNLF2a* through direct binding to proximal promoter ZREs

### 5.1 Introduction

Transcriptional regulation of genes is dependent on transcription factors (TFs) binding to cis-regulatory elements, and through specific protein-protein interactions, that lead to the recruitment of the basal transcriptional machinery at the gene core promoter within an accessible chromatin environment. Thus, promoters, the regions immediately upstream of the transcriptional start site of genes (TSS), have multiple adjacent TFs binding motifs which are commonly working synergistically through protein interactions (reviewed in (Lelli et al., 2012)).

As discussed in **Chapter 1**, Zta along with the other EBV important immediate-early protein Rta (BRLF1) are the master regulator of the events taking place during lytic cycle (Feederle et al., 2000; Miller et al., 2007). In particular, Zta is a viral site- specific transcription factor which is indispensable for inducing a complete EBV lytic cycle as well as initiating lytic viral DNA replication (Askovic and Baumann, 1997; Countryman and Miller, 1985; Feederle et al., 2000; Fixman et al., 1995; Grogan et al., 1987; Sarisky et al., 1996; Schepers et al., 1993a). Zta mediates the lytic cycle by inducing the expression of Rta (Sinclair et al., 1991), autoregulating its own expression (Flemington and Speck, 1990a; Lieberman and Berk, 1990), and inducing various other EBV lytic cycle genes (Chevallier-Greco et al., 1986).

Zta regulates its target genes by binding through a degenerate 7 bp motif (ZRE). 32 variants of this motif have been identified including the canonical AP-1 site (Chang et al., 1990; Farrell et al., 1989; Flower et al., 2011; Lieberman et al., 1990). These ZREs are present across the viral and cellular genome; in the EBV genome, ZREs can be identified in most of lytic cycle promoters within 500 bp of the TSS (Flower et al., 2011). In addition, Zta forms a highly stable TFIID-A complex, which is important to stabilise the TFIID binding at the core promoter site (Lieberman, 1994; Lieberman and Berk, 1991; Lieberman et al., 1997).



In recent years, not only Zta genome-wide binding to the viral and the cellular DNA was investigated (Bergbauer et al., 2010; Flower et al., 2011; Ramasubramanyan et al., 2012a; Ramasubramanyan et al., 2015) but Zta induced-changes in global cellular gene expression were analysed (Ramasubramanyan et al., 2015). This provided an overall view of Zta targeted genes. However, there is a lack of a complete understanding of Zta transcription mechanisms as only a small number of the viral and cellular promoters known to be responsive to Zta have been thoroughly characterised. For instance, as previously discussed in **Section 1.4**, the most widely investigated viral promoters are the *BZLF1* (Zp) (Flemington and Speck, 1990a) and *BRLF1* (Rp) promoters (Sinclair et al., 1991). In addition, other viral lytic promoters that have been investigated are *BMRF1*, *BHLF1* and *BALF2* (Holley-Guthrie et al., 1990; Hung and Liu, 1999; Kenney et al., 1989; Lieberman et al., 1989).

Investigating these promoters was very useful not only to understand Zta transcriptional regulation but EBV lytic cycle, particularly in EBV lytic semi-permissive cells (epithelial cells). For example, as described in **Section 1.4.4**, recent reports described the mechanism of Zta and Rta synergistic activation which was shown to be mediated by a cellular factor known as MCAF1 at *BHLF1* promoter (Chang et al., 2010). Another cellular protein Ku80, which is known to interact with Zta (Bailey et al., 2009a; Wiedmer et al., 2008), was shown to enhance the activation of *BHLF1* promoter (Chen et al., 2011). On the other hand, a few promoters were identified to be highly active or differentially expressed in a cell-specific manner (Becker et al., 1991; Bogedain et al., 1994; Lagenaur and Palefsky, 1999; Lemon et al., 1977; Lieberman et al., 1989; Young et al., 1991c). For instance, *BMRF1* promoter was shown to require Zta and Rta to give maximal activity in B cells while in epithelial cells Rta was dispensable (Holley-Guthrie et al., 1990). Another promoter which was shown to be highly expressed in oral hairy leukoplakia (OHL) is *BMRF2*. Unlike *BMRF1*, this promoter is not responsive to Zta but was induced by the phorbol ester TPA and characterised by the presence of GC box and TATA element. This promoter shows high basal levels in epithelial cells correlating with its high abundance in OHL (Lagenaur and Palefsky, 1999). A more detailed description of Zta responsive promoters is provided in **Section 1.4.4**.

In light of these studies, in this chapter, I show the analysis of a novel early- lytic promoter target for Zta. This promoter mediates the expression of BNLF2a, a small viral protein that plays a vital role in immune evasion, by targeting the transporter associated with antigen processing complex (TAP) (Hislop et al., 2007; Horst et al., 2009; Rensing et al., 2008; Wycisk et al., 2011). The role of BNLF2a in immune evasion is discussed in more details in **Chapter 1, Section 1.5.1**.

*BNLF2a* promoter (also known as ED-L2) is of great interest not only because it mediates the transcription of an essential EBV protein to avoid any immune response during lytic cycle (Croft et al., 2009; Quinn et al., 2014) but it was also described to have constitutive expression in stratified squamous epithelium of the tongue and esophagus (Jenkins et al., 1997; Nakagawa et al., 1997a; Nakagawa et al., 1997b; Wilson et al., 1990). It has also been employed in transgenic mice models to drive the specific expression of the oncogene Cyclin D1 and labelling fusion protein histone 2B-GFP in the esophagus cells, in an attempt to study the biology and cancer development in the esophagus (Nakagawa et al., 1997b; Roth et al., 2012).

Using luciferase reporter assay, a series of mutant constructs to dissect Zta response elements within the promoter were generated to investigate Zta regulation of this promoter in B cells as well as epithelial cells. A further mutation analysis of the promoter to investigate other potential cellular binding sites and Zta regulation mechanisms will be described in the next chapter.

## 5.2 Results

### 5.2.1 Zta binds at the *BNLF2a* promoter during lytic cycle

The ChIP-Seq data (Ramasubramanyan et al., 2012a) revealed two Zta binding peaks upstream *BNLF2a* promoter in Burkitt's lymphoma Akata cells as shown in **Figure 5-1 (A)**. Further *in silico* analysis of the DNA sequence under the binding peaks unfolded five Zta response elements. A large Zta binding peak is seen in very close proximity to the transcription start site (TSS) **Figure 5-1 (A)**. Under this large peak there is a cluster of three ZREs separated by very similar short distances. Around 80 bp separates ZRE1 and ZRE2, while 67 bp separates ZRE2 and ZRE3. On the other hand, around 580 bp away from *BNLF2a* transcription start site there is another smaller Zta binding peak above another two predicted ZREs within *BNLF2a* upstream sequence.

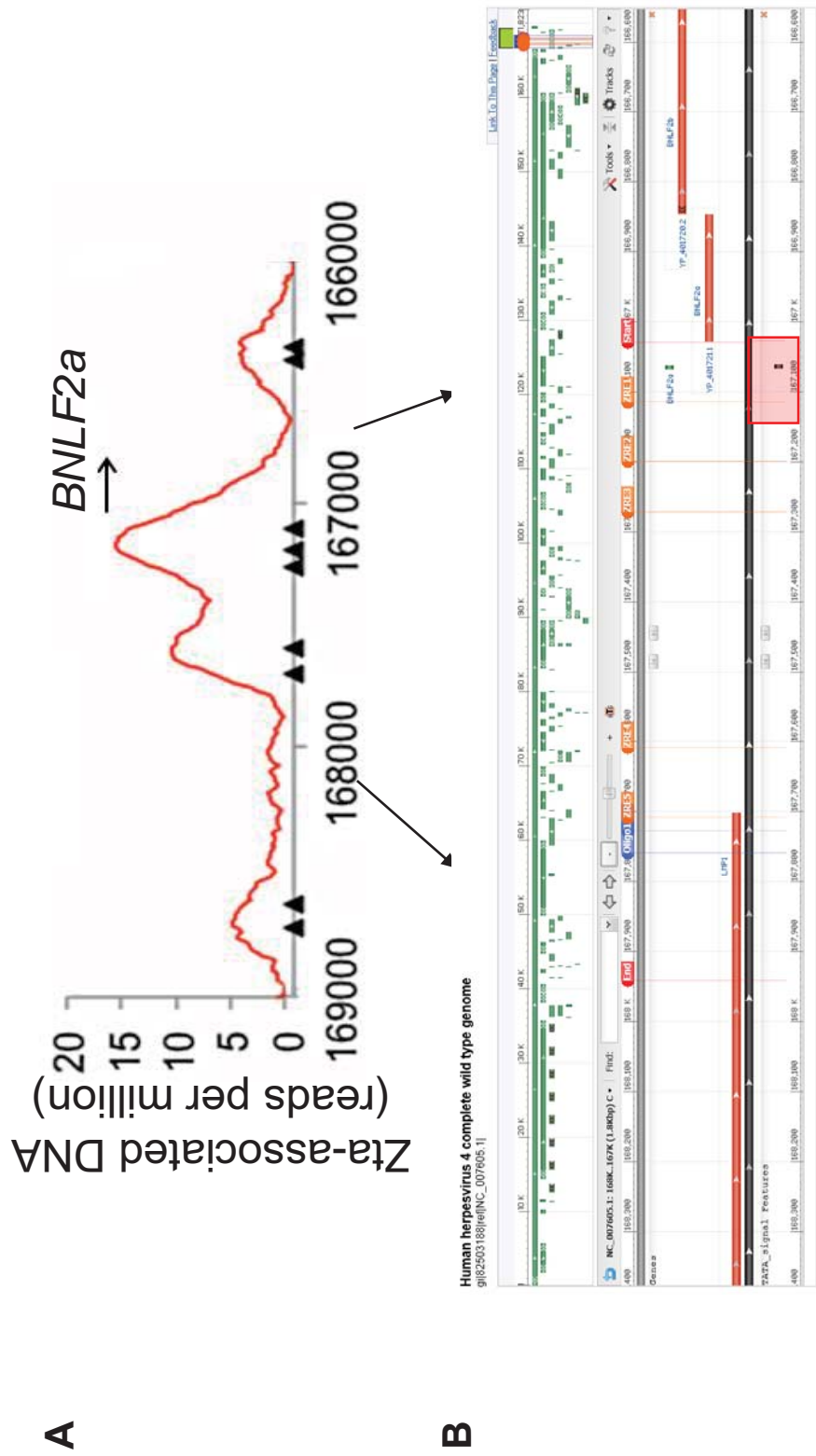
**Figure 5-1 (B)** represents a screen shot of a genome view of EBV wildtype strain (NC\_007605.1). This view shows the EBV genome region of *BNLF2a* gene, *BNLF2a* upstream (promoter) region investigated in this project, and genes in close proximity (*LMP1* and *BNLF2b*). The TATA box motif of *BNLF2a* is also highlighted.

The start and end of the promoter region (region between 167029-167941 bp) used in *BNLF2a* promoter luciferase reporter assay, which is discussed thoroughly in the next few sections, is shown between red flags in **Figure 5-1 (B)**. The exact location of each ZRE is also marked. The ZREs were given numbers from 1-5 where the closest one to the transcription site is ZRE1. The location of the primers used in Zta ChIP-qPCR experiment is also highlighted **Figure 5-1 (B)**. The coordinates of these oligonucleotides are also given in **Table 2-4**.

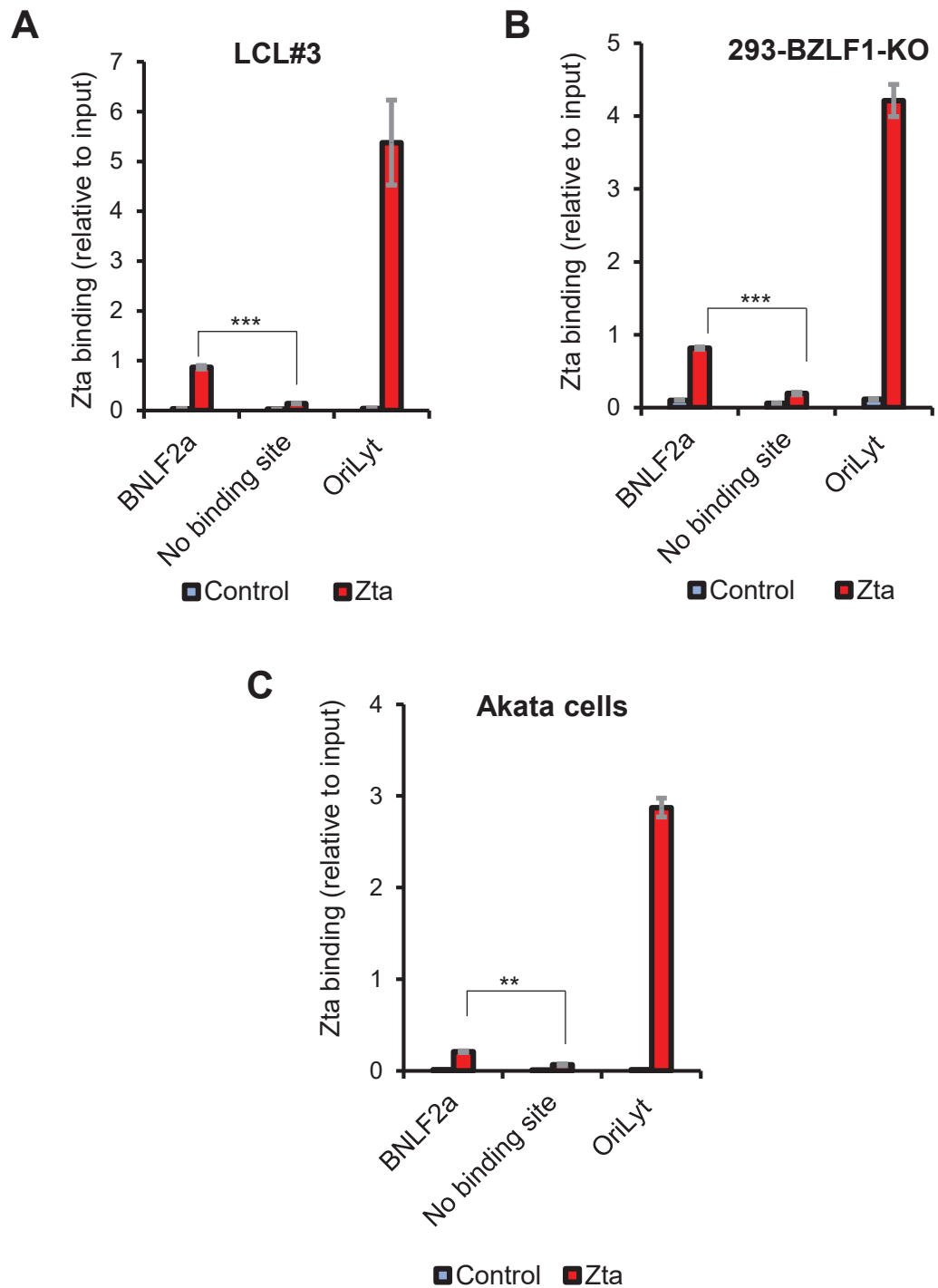
As previously shown in **Chapter 3**, we also looked into Zta binding at *BNLF2a* promoter in different EBV cell lines using ChIP-qPCR, particularly, the EBV non-B cell model available to us (293-BZLF1-KO cells). The ChIP-qPCR data is presented again in this chapter (**Figure 5-2 (A & B)**) along with ChIP-qPCR results in Akata cells (**C**).

Zta association with *BNLF2a* promoter region was assessed. The primers used in this qPCR are specific to a region near the second cluster of ZREs within the promoter region (**Figure 5-1 (B)**). Statistically significant enrichment for Zta at *BNLF2a* promoter was observed in all the investigated cell lines (P-values < 0.01 compared to the no binding control region).

The ChIP-qPCR data for Zta binding at *BNLF2a* promoter presented in **Figure 5-1** fall in agreement with the Zta ChIP-Seq data in lytic induced Akata B cells presented in **Figure 5-1 (A)**. Moreover, five ZRE motifs were identified within around 700 bp upstream of *BNLF2a* (**Figure 5-1 (B)**). These findings show preliminary evidence of a possible role for Zta in regulating *BNLF2a*, which is essential for the virus to avoid any host immune response during lytic cycle.



**Figure 5-1 Zta binding at *BNLF2a* promoter.** (A) ChIP-seq data shows two Zta binding peaks upstream of *BNLF2a* in Akata cells (Ramasubramanyan et al., 2012). Predicted ZREs are indicated by small black triangles. (B) A genome view of *BNLF2a* upstream region. *BNLF2a* and other viral genes in close proximity are shown. *In silico* predicted ZREs locations are marked (orange flags). TATA box motif is shown as a small black vertical bar (red box). The ChIP-qPCR primers used to validate Zta binding at this region are marked with blue flags. A 913 bp sequence including the five ZREs between 167029-167941 based on EBV wildtype strain (NC\_007605.1) was cloned in a reporter assay plasmid to analyse *BNLF2a* promoter activity (red flags) as will be explained in the following sections.



**Figure 5-2. Zta binding at BNLF2a promoter region in three different cell lines.** ChIP-qPCR in different EBV systems: Akata Burkitt's lymphoma cells induced into EBV lytic cycle by cross linking surface IgG, LCL3 cells which are spontaneously lytic lymphoblastoid cells and 293-BZLF1-KO cells which are HEK 293 cells stably infected with recombinant EBV containing *BZLF1* gene knockout. The qPCR assays were performed using primers for lytic origin of replication (OriLyt), non-binding region (OriLyt-Flank) and BNLF2a promoter. Values are presented as the mean of triplicate readings  $\pm$  SD. (\*\*\*) indicates  $p < 0.001$  while (\*\*) indicates  $p < 0.01$  compared to No binding site.

### 5.2.2 *BNLF2a* predicted ZRE sequences.

Given Zta binding at ZREs within the *BNLF2a* promoter, we hypothesised that this gene is a direct target for Zta. Thus, we wanted to investigate any functional relevance to this binding and to shed light on the possible mechanisms for this regulation.

We initially examined the predicted ZREs motif within the promoter. This is very important to determine what class of ZREs are present (unmethylated or methylated dependent ZREs) and whether any of these ZRE motifs were shown to bind Zta previously or at least have a role in regulating other viral promoters.

In **Figure 5-3**, the sequences and features of *BNLF2a*-predicated ZREs are listed as a table. ZRE1, which is closest to the TSS and separated by 41 nucleotides from the TATA box, has a CpG motif and is classified as type II ZRE. This means that Zta binds better when a methyl group is added to the cytosine nucleotide, but methylation is not essential for Zta binding (Flower et al., 2011; Karlsson et al., 2008b). The same motif was also described before in the EBV *BRLF1* promoter (Rp) (Bhende et al., 2004; Dickerson et al., 2009). The other four ZREs are class I (Karlsson et al., 2008b), which does not contain a CpG motif, and were described before in either other EBV promoters or within the OriLyt region. ZRE2 is a canonical AP-1 site (Lopez et al., 1993) and is found in EBV OriLyt (Lieberman et al., 1990). ZRE3 (TGTGCCA) and ZRE4 (TGGCACA) are a reverse complementary of each other. An identical motif to TGTGCCA is found in the EBV *BMRF1* promoter (Flower et al., 2011). ZRE5 is another known AP-1 site and has been described in various cellular promoters (Rani et al., 2009; Thierry et al., 1992) as well as in the EBV lytic *BFRF1* gene (Granato et al., 2006).

Given Name	5'-3' Sequence	EBV (NC_007605.1) Coordinates	Frequency of site in EBV genome	Description
<b>ZRE1</b>	T <u>CG</u> CTCA	167109-167115	17	ZRE1 is class II (contains CpG motif) (1&2). This ZRE motif is the same with ZRE2 in <i>BRLF1</i> promoter (Rp) (3 &4).
<b>ZRE2</b>	TGACACA	167194-167199	31	ZRE2 is the canonical AP-1 site (5) and has no CpG motif (class I). Also, found in EBV OriLyt (6)
<b>ZRE3</b>	TGTGCCA	167266-167272	41	ZRE3 is class I and found in BMRF1 promoter (1).
<b>ZRE4</b>	TGGCACA	167609-167603	41	ZRE4 is the reverse complementary sequence of ZRE3 found in BMRF1 promoter (1).
<b>ZRE5</b>	TGACTAA	167702-167708	8	ZRE5 is class I and described as AP-1 binding site in EBV BFRF1 (7) as well as various viral and cellular promoters (8&9)

**Figure 5-3 Zta response elements (ZREs) upstream of *BNLF2a* gene.** The table shows the predicted ZREs within 1 kb upstream of *BNLF2a* along with their genome coordinates according to the EBV wildtype strain (NC\_007605.1). Each ZRE location is marked (orange flag) in Figure 5-1. Brief description of each individual ZRE is also given. (1) (Flower et al., 2011), (2) (Karlsson et al., 2008), (3) (Bhende et al., 2004), (4) (Dickerson et al., 2009), (5) (Lopez et al., 1993), (6) (Lieberman et al., 1990), (7) (Granato et al., 2006) (8) (Rani et al., 2009), (9) (Thierry et al., 1992).



### 5.2.3 *BNLF2a* promoter constructs to investigate each ZRE function

We investigated the functional relevance of the identified ZREs using luciferase reporter assay. A 913 bp fragment immediately upstream of the start codon of *BNLF2a* that included the TATA box and all five ZREs was analysed in our reporter assay. Also, it is worth noting that the last part of this fragment overlaps with part of *LMP1* (3') end, as shown in **Figure 5-1 (B)**.

The *BNLF2a* promoter sequence (DNA fragment illustrated in **Figure 5-1 (B)**) was obtained as synthetic dsDNA fragment and was subcloned into the luciferase reporter assay plasmid pCpGL-basic upstream of the luciferase gene between BamHI and HindIII restriction enzyme sites, as demonstrated in the schematic diagram in **Figure 5-4**.

Mutant ZRE promoter variants were similarly subcloned to generate 12 mutant luciferase constructs, the details of which are provided later in **Figure 5-7**.

Each luciferase construct was co-transfected with either a control vector (pcDNA3) or His-Zta expression vector into EBV-negative Burkitt's lymphoma cell line (DG75) and/or human cervical carcinoma epithelial cells (HeLa cells). This approach allowed the comparison of luciferase activity in the presence and absence of Zta as demonstrated in **Figure 5-4 (B)**.

**Figure 5-5** represents the full DNA sequence of the wildtype *BNLF2a* promoter DNA insert (5'-3') used in our luciferase reporter analysis. Each predicted-ZRE sequence is highlighted in the figure along with the canonical TATA motif, which is 30 bp away from the transcription start site (TSS).

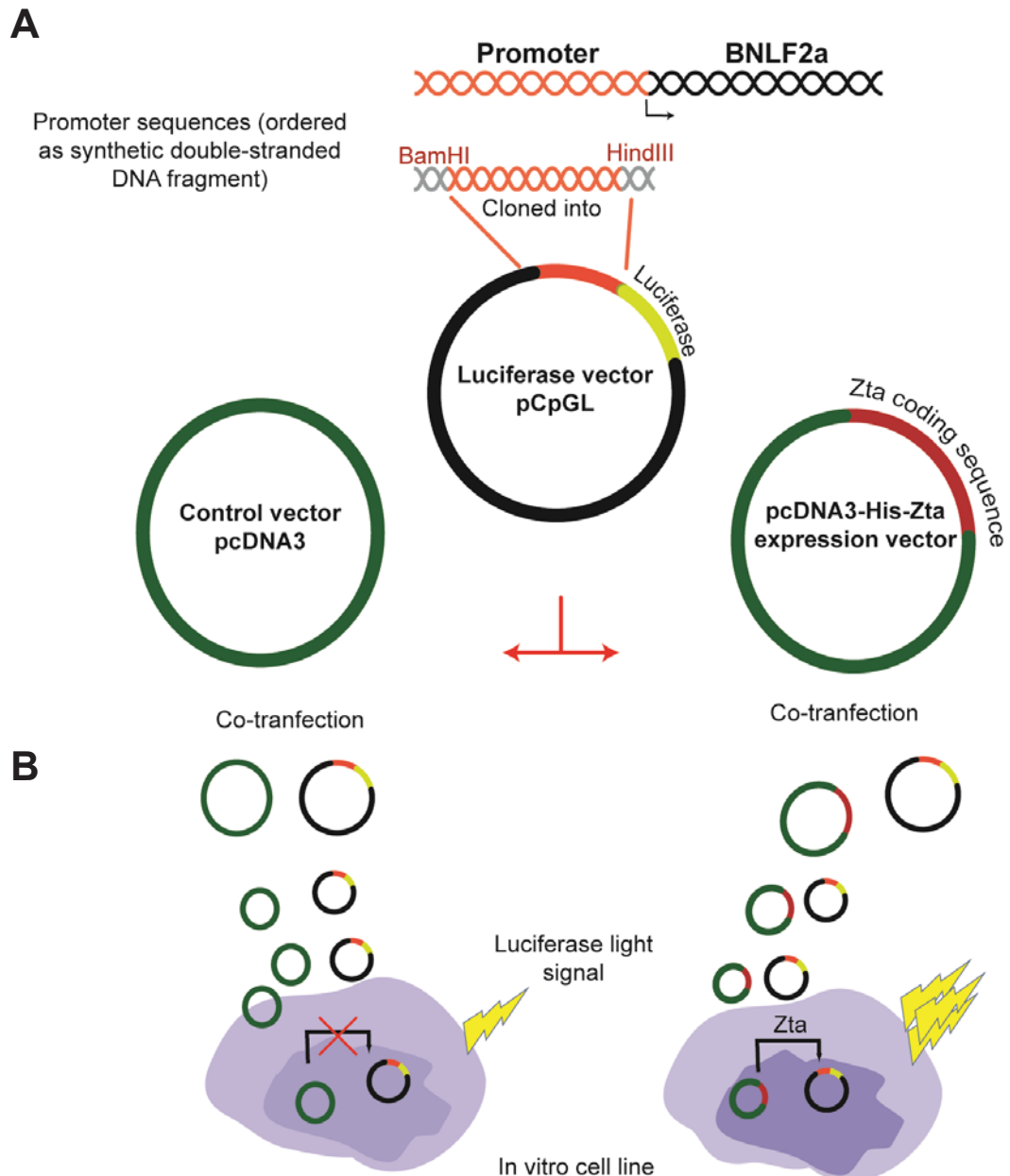
Furthermore, in **Figure 5-6 (A)**, a schematic cartoon of the promoter reporter constructs showing the sequence and location of each ZRE in relative to the gene start is given. Also, these locations are listed in a table (**Figure 5-6 (B)**).

As described earlier, the main aim was to map the ZRE motifs within the promoter that would possibly contribute to the promoter activity. To achieve

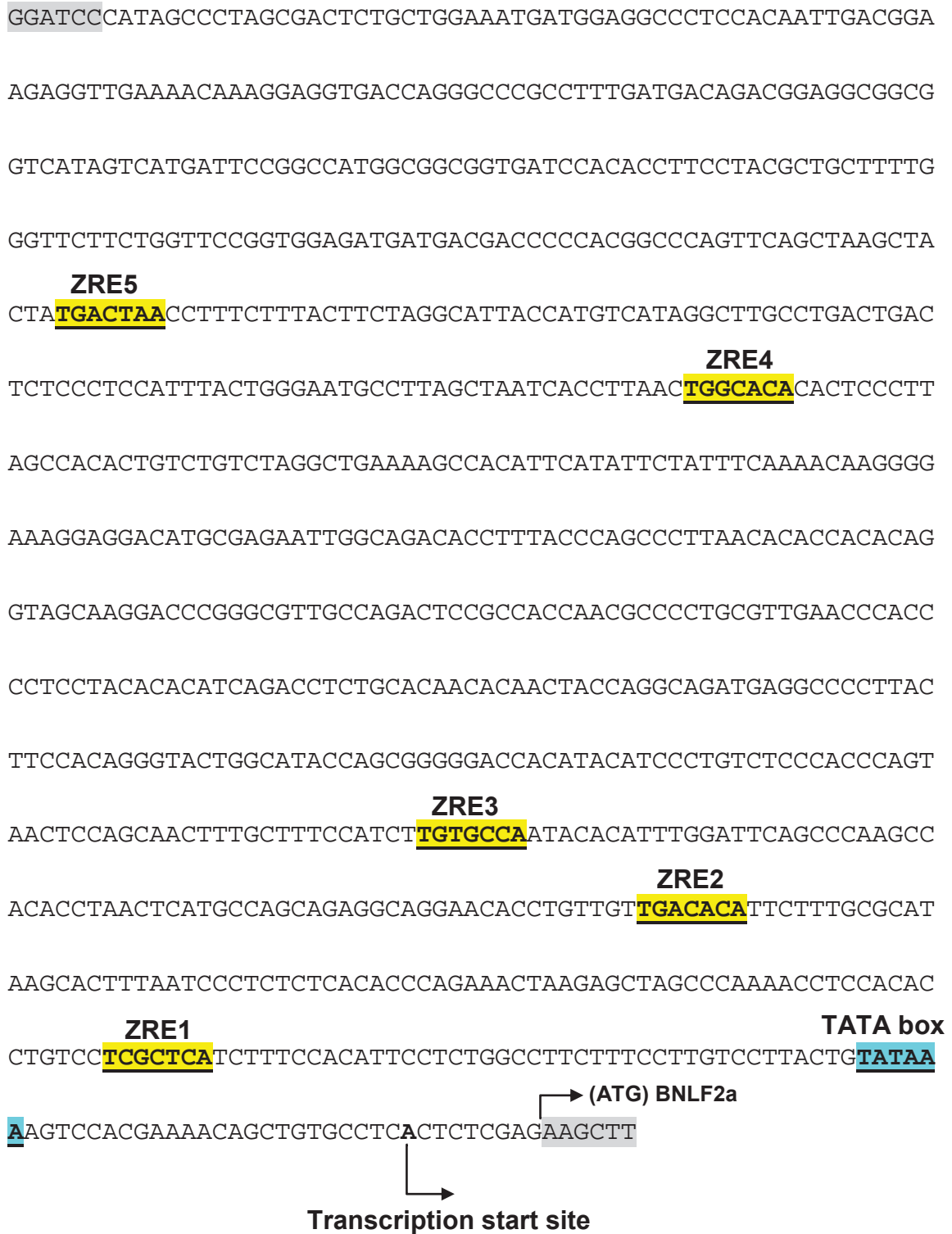
this, we generated a series of mutant promoter luciferase constructs as explained in the experimental design illustration in **Figure 5-4**. In each construct, either one or a different combination of ZREs were replaced each time by a random mutant motif (**CCCCTTT**).

The rationale of each mutant construct will be given in the relevant results section. Briefly, **Figure 5-7** provides a list of these mutant constructs consisting of the name given to the construct, a brief description of the mutated sites, and a small schematic for each construct that will be presented again in the relevant results section.

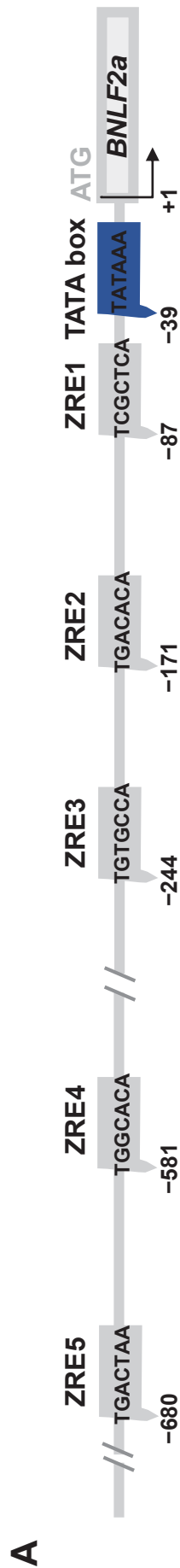
Our luciferase reporter assay approach allowed a simple and rapid analysis of *BNLF2a* promoter. This was facilitated by the technological advances in DNA fragment synthesis which enabled us to generate numerous mutant promoter constructs to investigate Zta regulation of *BNLF2a*.



**Figure 5-4 Investigating *BNLF2a* promoter activity.** A schematic representation of the luciferase reporter assay experiment carried out to investigate *BNLF2a* transcriptional regulation. **(A)** A synthetic DNA fragment (orange) spanning around 900 bp upstream of *BNLF2a* gene start was cloned between BamHI and HindIII restriction enzyme sites in pCpGL-basic luciferase vector (black plasmid). Additionally, 20 mutant *BNLF2a* promoter variants were cloned in pCpGL vector. **(B)** Each promoter variant luciferase vector was co-transfected with either a control vector (pcDNA3) (green plasmid) or His-Zta expression vector (green and red plasmid) in EBV-negative Burkitt's lymphoma cell line (DG75) and/or epithelial cells (HeLa cells). Luciferase signal was compared in the absence and presence of Zta.



**Figure 5-5 *BNLF2a* promoter sequence along with the predicted ZRE motifs.** A schematic illustration of the sequence used in *BNLF2a* promoter constructs. A 913 bp sequence with the genome coordinates 167029-167941 based on EBV wildtype strain (NC\_007605.1) was used to analyse *BNLF2a* promoter activity. Sequence analysis revealed five Zta response elements (ZREs) (highlighted in yellow) distributed within 700 bp upstream of *BNLF2a* gene. A canonical TATA box motif (highlighted blue) is located 30 bp upstream of the transcription start site. The gene start and transcription start site (TSS) are indicated with arrows. The restriction enzyme sites used in cloning are highlighted in grey.









**B**

Given Name	5'-3' Sequence	EBV (NC_007605.1) coordinates (relative to the first position of the start codon (+1))
ZRE1	TCGCTCA	-87 to -81
ZRE2	TGACACA	-171 to -165
ZRE3	TGTGCCA	-244 to -238
ZRE4	TGGCACA	-581 to -575
ZRE5	TGACTAA	-680 to -674

**Figure 5-6 A schematic diagram of *BNLf2a* promoter.** A 913 bp DNA fragment between 167029-167941 based on EBV wildtype strain (NC\_007605.1) was used to analyse *BNLf2a* promoter activity. **(A)** A schematic illustration of *BNLf2a* promoter. ZREs (grey boxes) are given numbers from 1-5 where ZRE1 is the closest to the gene start. TATA box motif is shown (dark blue box). The position of each ZRE (relative to the gene start) is shown. **(B)** The sequence and position of each ZRE relative to the gene start are listed in a table.

Promoter construct name	Description	Schematic diagram
BNLF2a 1-5	5 ZREs left intact	
BNLF2a Null	None of the ZREs left intact	
BNLF2a 1-3	ZREs 1-3 left intact	
BNLF2a 4-5	ZREs 4-5 left intact	
BNLF2a 2-5	ZREs 2-5 left intact	
BNLF2a 1&3-5	ZREs 1 and 3-5 left intact	
BNLF2a 1-2&4-5	ZREs 1-2&4-5 left intact	

Promoter construct Name	Description	Schematic diagram
BNLF2a 1	Only ZRE1 left intact	
BNLF2a 2	Only ZRE2 left intact	
BNLF2a 3	Only ZRE3 left intact	
BNLF2a 1&2	ZREs 1&2 left intact	
BNLF2a 1&3	ZREs 1&3 left intact	
BNLF2a 2&3	ZREs 2&3 left intact	

**Figure 5-7 *BNLF2a* ZREs luciferase promoter constructs.** All the luciferase reporter constructs generated to address ZREs contribution to promoter activity are listed. The construct given name, ZRE mutations and a schematic diagram of each constructs are given. To mutate each ZRE motif, the wildtype ZRE motif was replaced with a generic sequence (CCCCTTT).

#### 5.2.4 Zta activates *BNLF2a* through promoter ZREs in different cell lines

Various *BNLF2a* promoter reporter constructs were generated to investigate Zta activation of *BNLF2a* promoter and to identify the individual ZRE/s responsible for this activation, as explained in the previous sections.

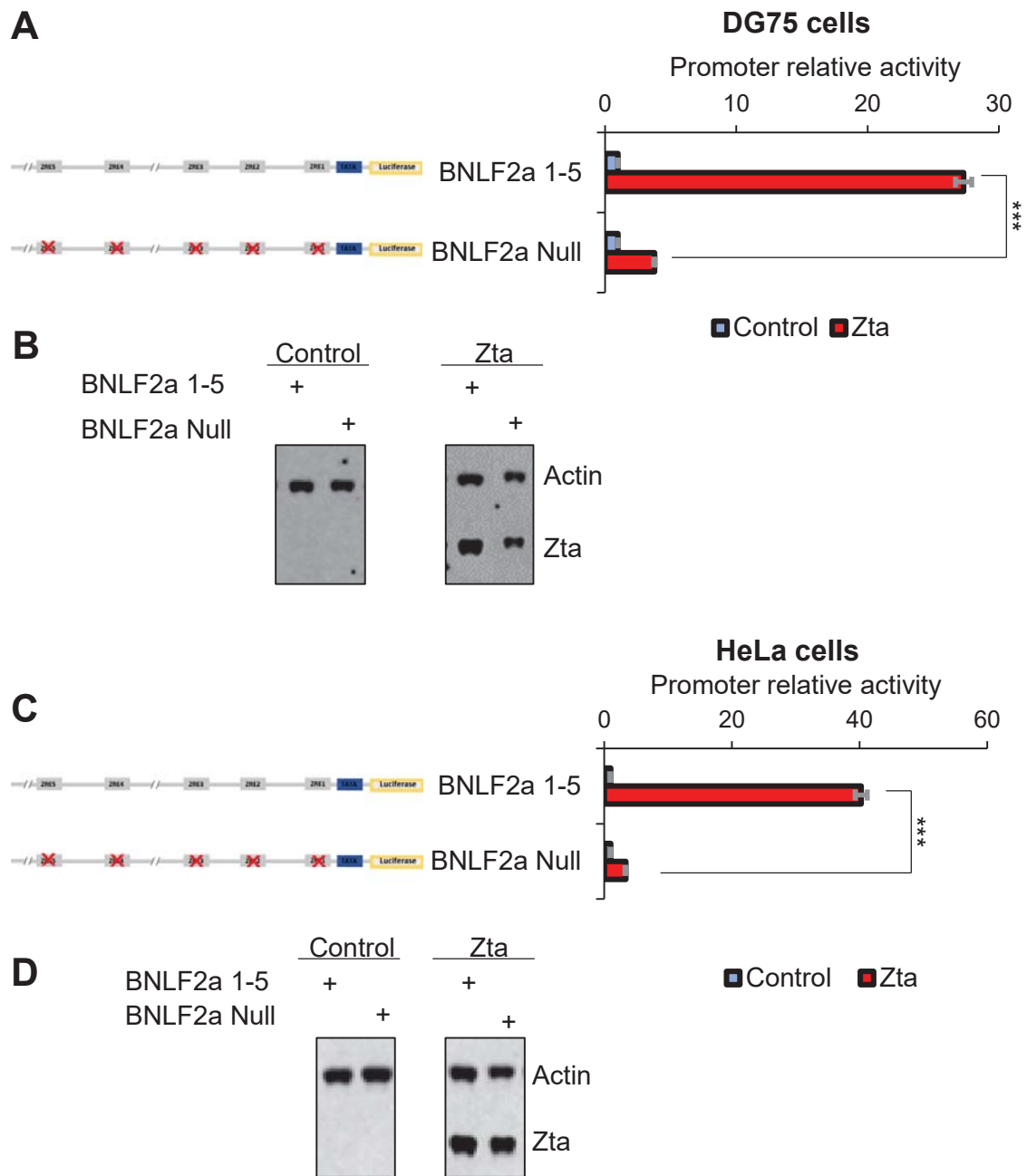
We initially investigated whether Zta alone is capable of activating *BNLF2a* and whether this activation is dependent on the direct binding of Zta through its response elements (ZREs). Thus, we compared the activity of wildtype *BNLF2a* promoter (BNLF2a 1-5) against a construct with mutated ZRE sites (BNLF2a Null). This comparison was carried out in the presence and absence of Zta expression in EBV-negative DG75 B-cells as well as HeLa cells. A representative data of the BNLF2a 1-5 and BNLF2a Null promoter luciferase relative activity is shown in **Figure 5-8**. The data is presented as fold change of the activity in the presence of Zta over the absence of Zta (pcDNA3 control sample).

The relative activity of the wildtype promoter construct (BNLF2a 1-5) was at least 7-fold greater than BNLF2a Null in which all the predicted ZRE sites were mutated (**Figure 5-8 (A and C)**). This finding was consistent between EBV-negative DG75 cells (**A**) and HeLa cells (**C**). Using student t-test, a statistically significant difference in promoter activity between BNLF2a 1-5 and BNLF2a Null construct in the presence of Zta was observed (P-value <0.001).

**Figure 5-8 (B and D)** represents a western blot analysis for actin (loading control) and Zta for each sample. Comparable levels of Zta expression can be seen for all samples while the absence of Zta is confirmed in the control samples. For each sample, the intensity of the actin band was used to normalise the luciferase relative light units value for that sample as previously described in **Chapter 2**.

Clearly, the data presented in **Figure 5-8**, strongly indicates that Zta, with no other viral proteins, activates *BNLF2a* promoter in both cell lines tested. This activation is dependent on the presence of intact ZREs, as the promoter activity was massively attenuated when these sites were mutated.





**Figure 5-8 Zta activates *BNLF2a* through ZREs.** The figure shows luciferase reporter assay analysis of the wildtype *BNLF2a* construct (*BNLF2a* 1-5) and mutated ZREs (*BNLF2a* Null) construct co-transfected with either Zta expression vector (red bars) or control vector (blue bars) in EBV-negative Burkitt's lymphoma cells (DG75 cells) (**A**) and epithelial cells (HeLa cells) (**C**). (**B** and **D**) Western blots show similar levels of Zta transfection in each cell line. Values are normalised to actin levels and presented as fold change in relative to the control for each sample. Error bars represent the mean of triplicate reads  $\pm$  SD. (\*\*\*) indicates  $p < 0.001$  compared to *BNLF2a* Null in the presence of Zta.

### 5.2.5 The proximal ZREs to the TSS are essential for *BNLF2a* activation.

Following our first intriguing observation where Zta activated the promoter through ZREs, the following logical questions were raised: Which ZREs are contributing to the promoter activity? Does the arrangement of the ZREs relate to their function?

To better understand these questions, we were equally interested in what appears to be two cluster regions of ZREs separated by around 330 bp. This was largely based on the observation that there were two Zta binding peaks in the ZRE regions ChIP-Seq data, as shown in **Figure 5-1**.

Thus, we generated two mutant *BNLF2a* promoter constructs containing either the first three proximal ZREs (ZRE1-3) (*BNLF2a* 1-3) or the last two distal ZREs (ZRE 4-5) (*BNLF2a* 4-5) in the wildtype form.

The promoter activity of *BNLF2a* 1-3 and *BNLF2a* 4-5 reporter constructs was investigated in a luciferase reporter assay experiment, as explained in the previous sections. This analysis was carried out in the presence and absence of Zta expression in EBV-negative DG75 B-cells as well as HeLa cells.

**Figure 5-9** shows a representative experiment, in which wildtype *BNLF2a* 1-5, *BNLF2a* Null, *BNLF2a* 1-3 and *BNLF2a* 4-5 constructs were compared. The data are presented as promoter relative activity (fold change) in the presence of Zta over the absence of Zta.

Interestingly, the *BNLF2a* 1-3 construct was still activated in the presence of Zta at levels that were very comparable to the levels seen in the wildtype *BNLF2a* 1-5 construct in both cell lines tested (**Figure 5-9 (A and C)**).

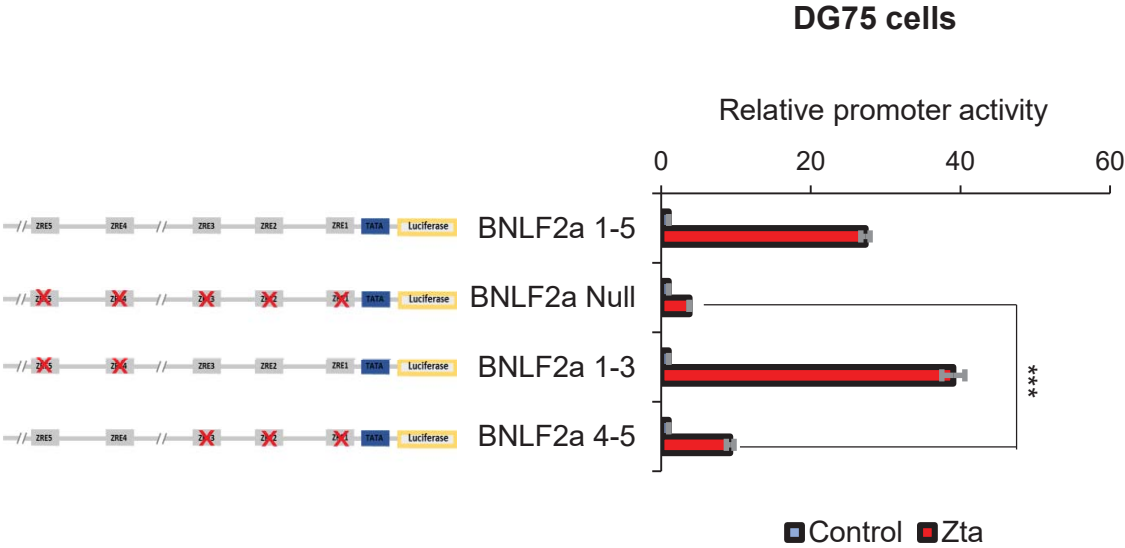
The relative promoter activity of *BNLF2a* 1-3 was at least 9-fold greater than *BNLF2a* Null. This finding was consistent between EBV-negative DG75 cells (**A**) and epithelial cells (HeLa cells) (**C**).

On the other hand, the relative promoter activity of BNL2a 4-5 construct, which included the intact two distal ZREs, showed a very reduced activity in DG75 cells and a completely insignificant change in HeLa cells, when compared to BNL2a Null.

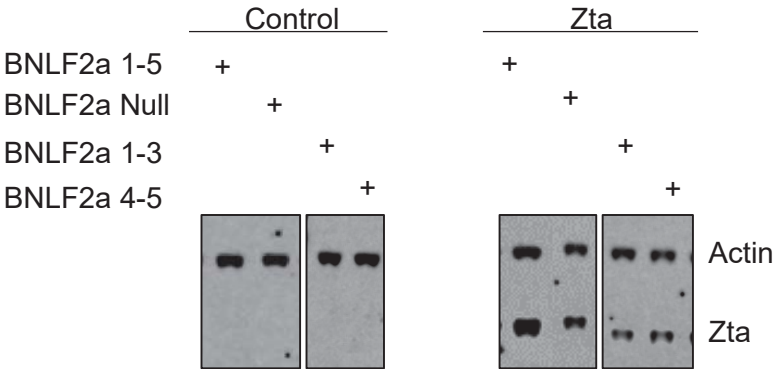
**Figure 5-9 (B and D)** represents a western blot analysis for actin (loading control) and Zta for each sample. Comparable levels of Zta expression can be seen for all samples while the absence of Zta is confirmed in the control samples. For each sample, the intensity of the actin band was used to normalise the luciferase relative light units value for that sample, as previously described in **Chapter 2**.

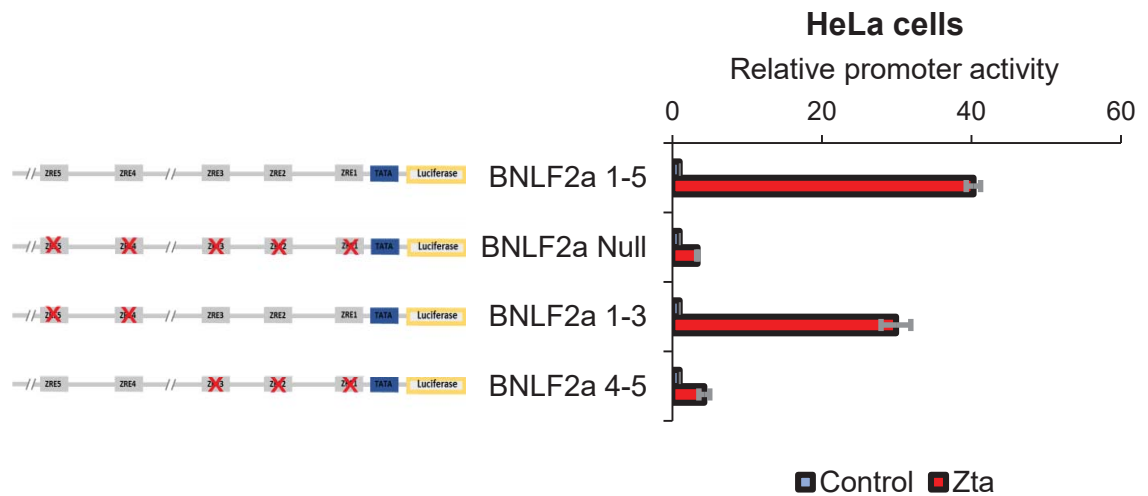
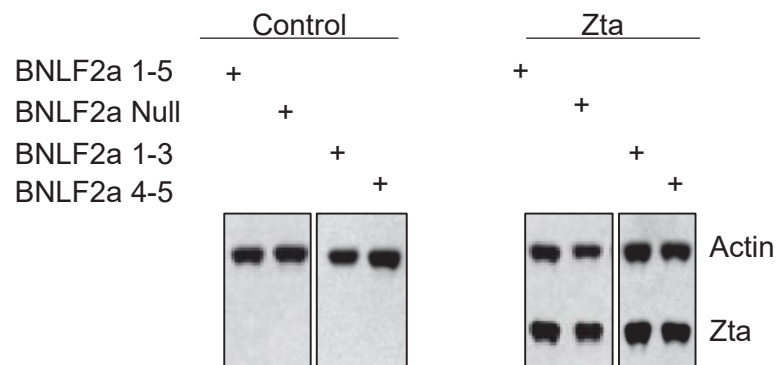
Clearly, the data presented in **Figure 5-9**, strongly indicate that *BNL2a* promoter response to Zta is largely driven from the first three ZREs (ZRE 1-3) promoter in both cell lines tested (DG75 and HeLa cells).

A



B



**C****D**

**Figure 5-9 The proximal ZREs contribute the most to *BNLF2a* activity.** The figure shows luciferase reporter analysis of the wildtype *BNLF2a* construct (*BNLF2a* 1-5), mutated ZREs (*BNLF2a* Null) construct as well as a construct with only the three proximal ZREs left (*BNLF2a* 1-3) and a construct with the two distal ZREs left (*BNLF2a* 4-5). *BNLF2a* promoter luciferase constructs were co-transfected with either Zta expression vector (red bars) or control vector (blue bars) in EBV-negative Burkitt's lymphoma cells (DG75 cells) (previous page) **(A)** and epithelial cells (HeLa cells) **(C)**. **(B and D)** Western blots show similar levels of Zta transfection in each cell line. Values are normalised to actin levels and presented as fold change in relative to the control for each sample. Error bars represent the mean of triplicate reads  $\pm$  SD. *BNLF2a* 4-5 is significantly different from *BNLF2a* Null ( $p < 0.001$ ) in DG75 cells but not in HeLa cells. (\*\*\*) indicates  $P < (0.001)$ .

### 5.2.6 ZREs functional redundancy

Based on our previous finding where the first three ZREs were shown to be critical for *BNLF2a* promoter response to Zta activation, we asked the question whether all of the three ZREs are needed, or is it possible to further identify the individual ZREs responsible for the promoter activity. To address this question, we used the same luciferase approach and generated additional mutant *BNLF2a* promoter constructs that accounted for all combinations of the first three ZREs.

Initially, we generated the three separate luciferase promoter constructs *BNLF2a* 2-5, *BNLF2a* 1&3-5 and *BNLF2a* 1-2&4-5, where each contained a mutant form of one ZRE out of the first three. This was to determine whether any single ZRE was essential for *BNLF2a* activation.

**Figure 5-10** shows a representative experiment in which the promoter activity of *BNLF2a* 1-5, *BNLF2a* Null, *BNLF2a* 2-5, *BNLF2a* 1&3-5 and *BNLF2a* 1-2&4-5 constructs were compared. The data is presented as fold change of the activity in the presence of Zta over the absence of Zta (control) in EBV-negative DG75 cells as well as HeLa cells.

The *BNLF2a* 2-5 construct, containing ZREs 2-5 and mutated ZRE1, showed promoter activity in the presence of Zta but at slightly reduced levels compared to the wildtype promoter construct *BNLF2a* 1-5 in both cell lines (EBV negative DG75 cells and HeLa cells) (**Figure 5-10 A and C, respectively**).

The *BNLF2a* 1-2&4-5 construct, containing ZREs 1-2&4-5 and mutated ZRE3, showed a similar level of activation compared to *BNLF2a* 2-5 in the presence of Zta in both cell lines (**Figure 5-10 A and C, respectively**).

On the other hand, the *BNLF2a* 1&3-5 construct, containing ZREs 1&3-5 and mutated ZRE2, was still largely activated in the presence of Zta, but to a much lower degree than the levels reported for *BNLF2a* 2-5 construct and *BNLF2a* 1-2&4-5 construct.

**Figure 5-10 (B and D)** represents a western blot analysis for actin (loading control) and Zta for each sample, as discussed in the previous figure. Comparable levels of Zta expression can be seen for all samples while the absence of Zta is confirmed in the control samples. For each sample, the intensity of the actin band was used to normalise the luciferase relative light units value for that sample, as previously described in **Chapter 2**.

The data presented in **Figure 5-10**, suggests that no single ZRE of the first proximal ZREs is key motif in Zta activation of *BNLF2a*.

This initial finding suggested a possible functional redundancy of ZREs but did not eliminate whether the wildtype activity of *BNLF2a* promoter requires one or more ZREs.

We therefore generated additional constructs to determine whether one or more ZRE were enough to induce the wildtype promoter activity. A list of all constructs is provided in **Figure 5-7**. Also, based on the fact that there were no major differences in the results between the two cell lines used in our experiments, a decision was made to proceed with transfecting the new constructs into HeLa cells only. Unlike DG75 cells, HeLa cells provided a much more consistent transfection efficiency as electroporation was not required.

**Figure 5-11 (A)**, represents data showing the relative promoter activity of the *BNLF2a* promoter constructs *BNLF2a 1*, *BNLF2a 2* and *BNLF2a 3*, which contains only one intact ZRE in HeLa cells only. Comparing the relative activity of these constructs to the wildtype promoter construct (*BNLF2a 1-5*) suggested that no single ZRE was capable of driving the wildtype promoter activity.

The *BNLF2a 3* construct containing intact ZRE3 only was not responsive to Zta induction, while *BNLF2a 1* and *BNLF2a 2* constructs containing intact ZRE1 and ZRE2 respectively showed a reduced but statistically significant activation when compared to *BNLF2a Null*. This activation was slightly more prominent in *BNLF2a 2*. These results are in agreement with the results presented in **Figure 5-10** where mutating one of the three proximal ZREs showed no adverse effects

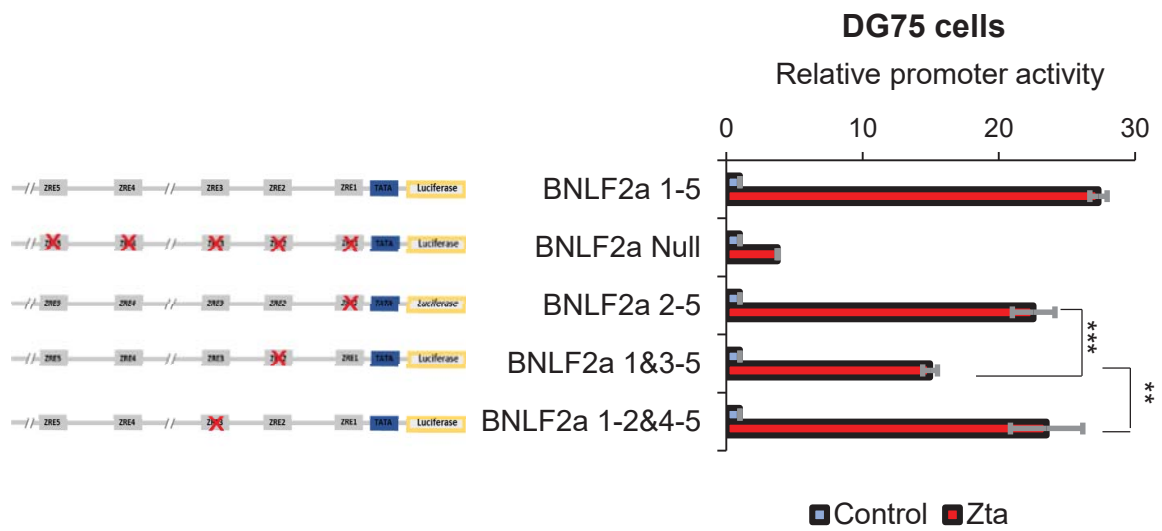
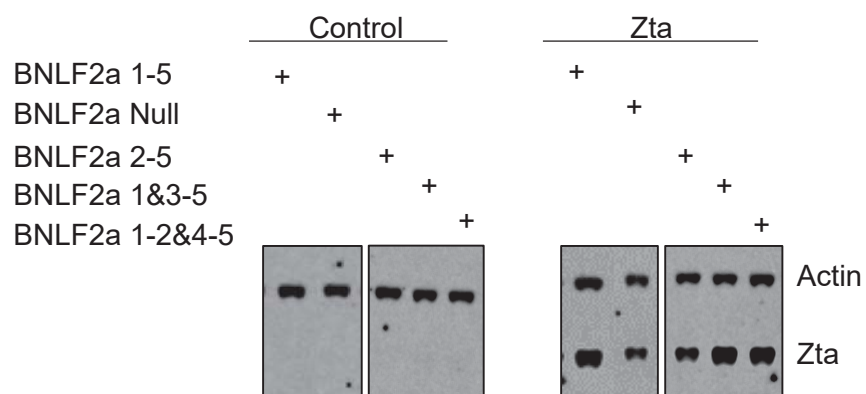
on the promoter activity. Also, ZRE2 had a more prominent effect, suggesting a potential dominant role for ZRE2.

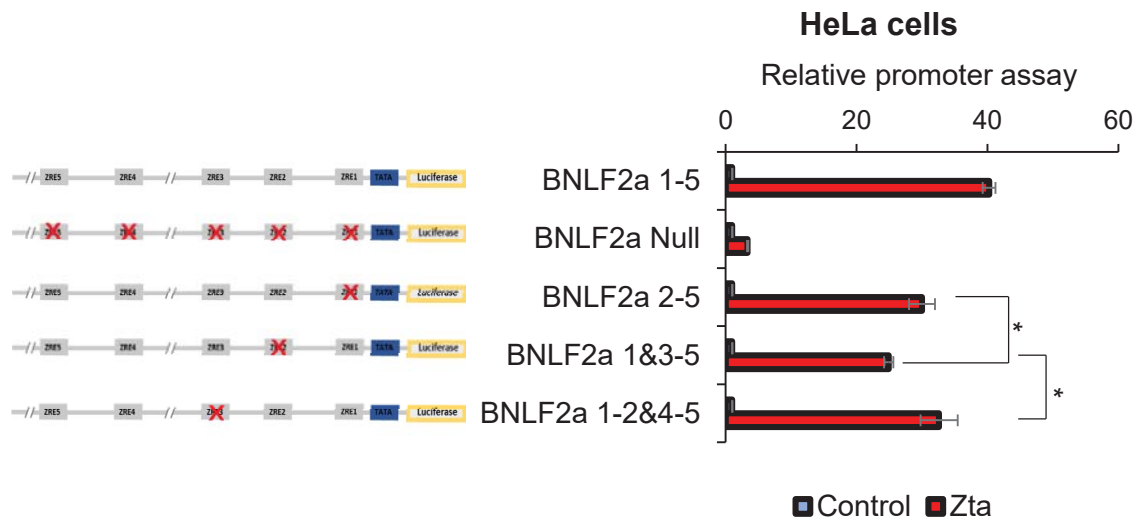
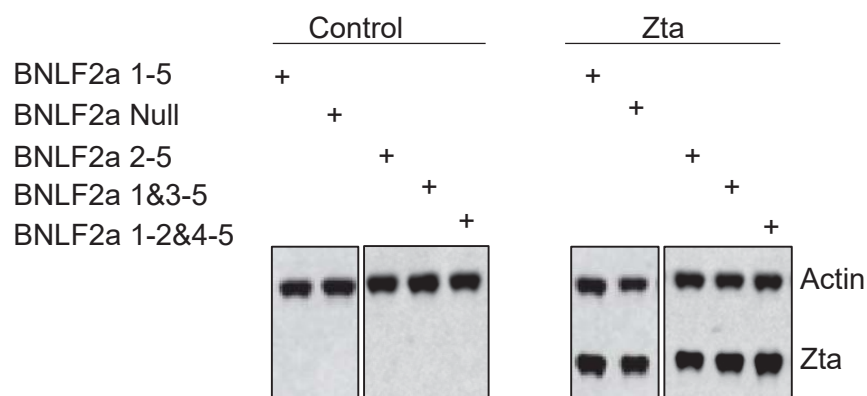
As previously discussed, **Figure 5-11 (B)** shows a western blot analysis of the levels of Zta transient expression along with actin (loading control).

We hypothesised that the wildtype promoter activity is very likely to be derived from at least two functional ZREs of the proximal ZREs based on the findings presented in **Figure 5-10** and **5-11**. To test this hypothesis, we additionally generated the *BNLF2a* promoter constructs BNL2a 1&2, BNL2a 1&3 and BNL2a 2&3. Each of these constructs contains two different intact ZREs of the first three proximal ZREs within *BNLF2a* promoter.

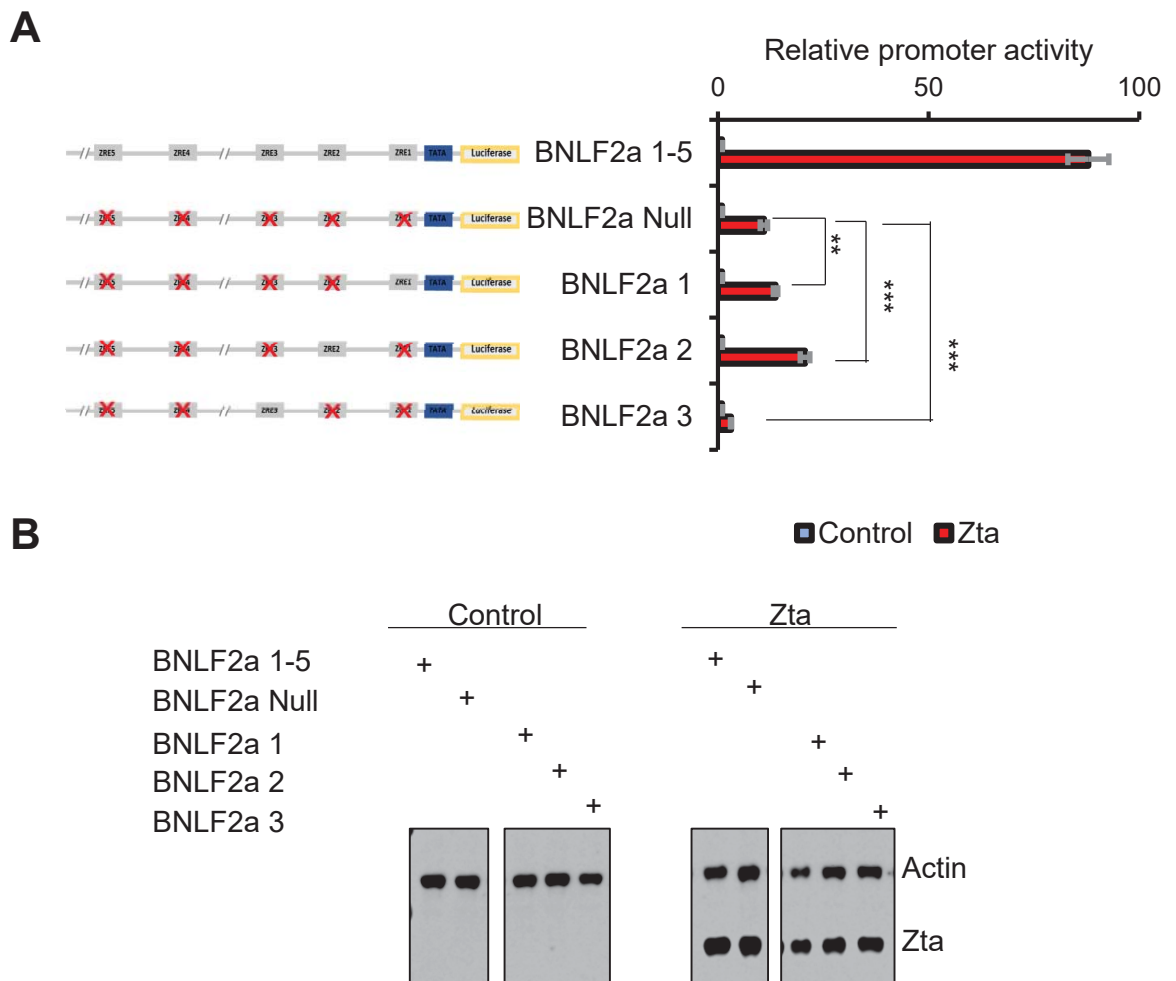
**Figure 5-12 (A)**, represents data demonstrating the promoter activity in the presence of Zta for BNL2a 1&2, BNL2a 1&3 and BNL2a 2&3 constructs. Interestingly, leaving two ZREs each time has restored the promoter activity to levels similar to the level of the wildtype promoter activity (BNL2a 1-5). These levels were at least 6- fold greater than BNL2a Null. Noticeably, the BNL2a 1&3 construct, containing intact ZRE1, ZRE3 and mutated ZRE2, resulted in a slightly significant reduced activation when compared to BNL2a 1&2 and BNL2a 2&3 constructs ( $P < 0.05$ ). This supports ZRE2 key role in the promoter activation. Furthermore, the results presented in **Figure 5-12** have majorly reproduced the findings in **Figure 5-10**. However, unlike in **Figure 5-10**, the constructs presented in **Figure 5-12** contained mutated ZRE4-5. This confirms our initial observation of a less important role for ZRE 4-5 in *BNLF2a* promoter activation, as shown in **Figure 5-9 (C)**.



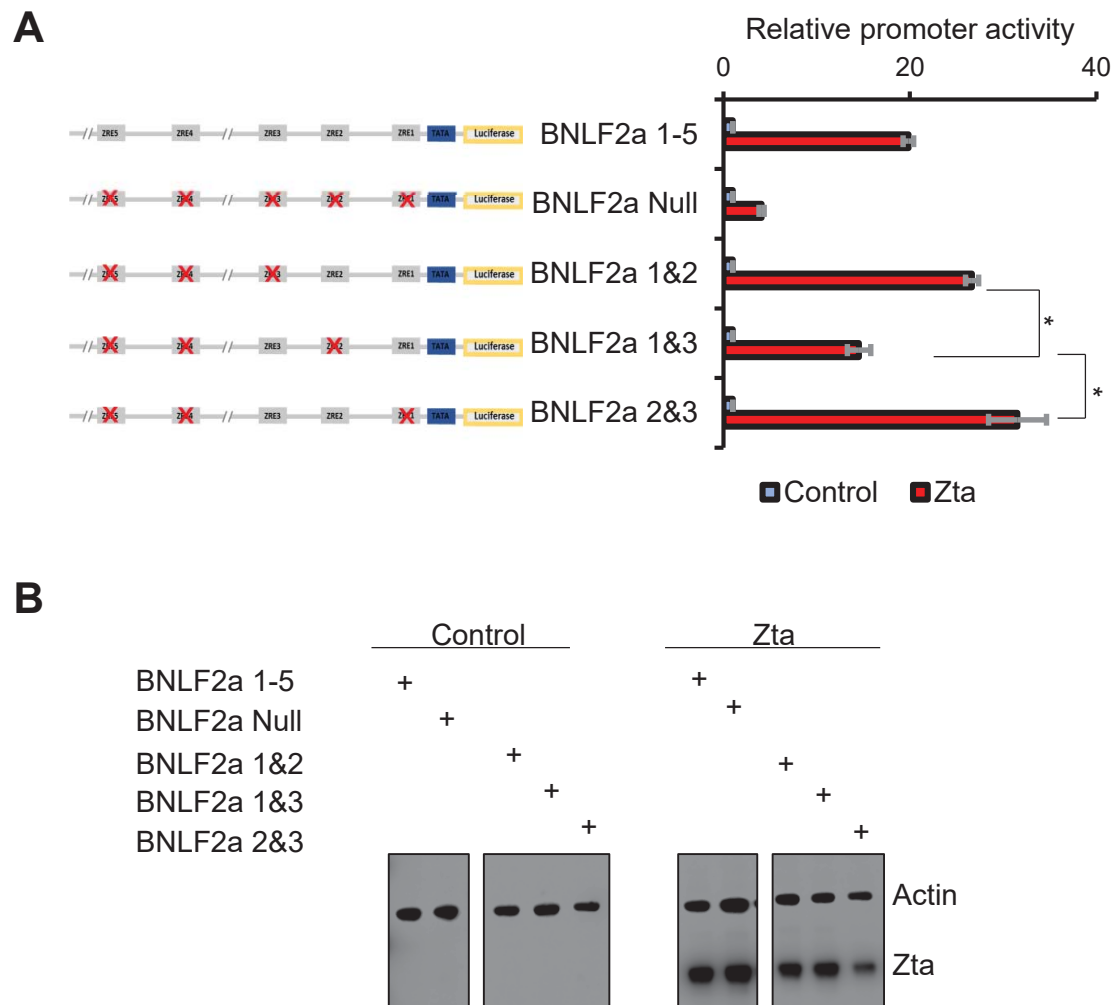
**A****B**

**C****D**

**Figure 5-10 Mutating one ZRE of the proximal ZREs has no profound effect on promoter activity.** The figure shows luciferase reporter analysis of the wildtype BNLf2a construct (BNLF2a 1-5), mutated ZREs (BNLF2a Null) construct as well as constructs with only one of the three proximal ZREs is mutated each time. BNLf2a promoter luciferase constructs were co-transfected with either Zta expression vector (red bars) or control vector (blue bars) in EBV-negative Burkitt's lymphoma cells (DG75 cells) (previous page) (A) and epithelial cells (HeLa cells) (C). (B and D) Western blots show similar levels of Zta transfection. Values are normalised to actin levels and presented as fold change in relative to the control for each sample. Error bars represent the mean of triplicate reads  $\pm$  SD. (\*\*\*) indicates  $p < 0.001$ , (\*\*) indicates  $p < 0.01$  and (\*) indicates  $P < (0.05)$  compared to BNLf2a 1&3-5 in the presence of Zta.



**Figure 5-11 The effect of leaving one ZRE of the proximal ZREs on promoter activity.** The figure shows luciferase analysis of the wildtype BNL2a construct (BNLF2a 1-5), mutated ZREs (BNLF2a Null) construct as well as constructs with only one of the three proximal ZREs remaining each time (**A**). *BNLF2a* promoter luciferase constructs were co-transfected with either Zta expression vector (red bars) or control vector (blue bars) in epithelial cells (HeLa cells). (**B**) Western blot shows similar levels of Zta transfection. Values are normalised to actin levels and presented as fold change in relative to the control for each sample. Error bars represent the mean of triplicate reads  $\pm$  SD. (\*\*\*) indicates  $p < 0.001$ , and (\*\*) indicates  $p < 0.01$  compared to BNL2a Null in the presence of Zta.



**Figure 5-12 The effect of leaving two ZRE of the proximal ZREs on promoter activity.** The figure shows luciferase analysis of the wildtype BNL2a construct (BNLF2a 1-5), mutated ZREs (BNLF2a Null) construct as well as constructs with only two of the three proximal ZREs remaining each time (**A**). *BNLF2a* promoter luciferase constructs were co-transfected with either Zta expression vector (red bars) or control vector (blue bars) in epithelial cells. (**B**) western blot shows similar levels of Zta transfection. Values are normalised to actin levels and presented as fold change in relative to the control for each sample. Error bars represent the mean of triplicate reads  $\pm$  SD. (\*) indicates  $P < (0.05)$  compared to BNL2a 1&3 in the presence of Zta.

### 5.2.7 *In vitro* analysis of Zta binding to *BNLF2a* ZREs

We further investigated Zta binding to each ZRE *in vitro* in an attempt to complete a simple model for Zta activation of *BNLF2a* promoter. To do this, we carried out electrophoresis mobility shift assay (EMSA) using fluorescently labelled probes of each *BNLF2a* ZRE, as previously described in **Chapter 2**. Furthermore, we generated a recombinant Zta expression vector and expressed this recombinant protein, which resembles a His-GFP tagged basic leucine zipper (bZIP) domain of Zta, in *E. coli*.

#### 5.2.7.1 bZIP Zta recombinant protein expression and purification

Prior to the EMSA, we needed to produce purified Zta protein to address Zta binding specificity to *BNLF2a* ZREs. As a part of a collaboration project with Oxford Protein Production Facility (OPPF), several recombinant expression plasmids were generated to test Zta expression in *E. coli* as well as insect cells. Upon initial expression screens in *E. coli*, a recombinant plasmid for a GFP-tagged Zta bZIP domain (DNA binding domain) showed promising protein expression. Therefore, this recombinant protein plasmid was selected to be expressed and purified at a large-scale and subsequently used in the EMSA experiment.

The methods used in generating this plasmid and expressing Zta recombinant protein are briefly outlined in **Figure 5-13** and **Figure 5-15**. By utilising the OPPF high-throughput methods, the plasmid was generated using In-Fusion ligation where the bZIP domain (168-245 aa) DNA sequence was subcloned into the pOPINN-GFP expression plasmid, which has N-terminus tag sequence for His-GFP.

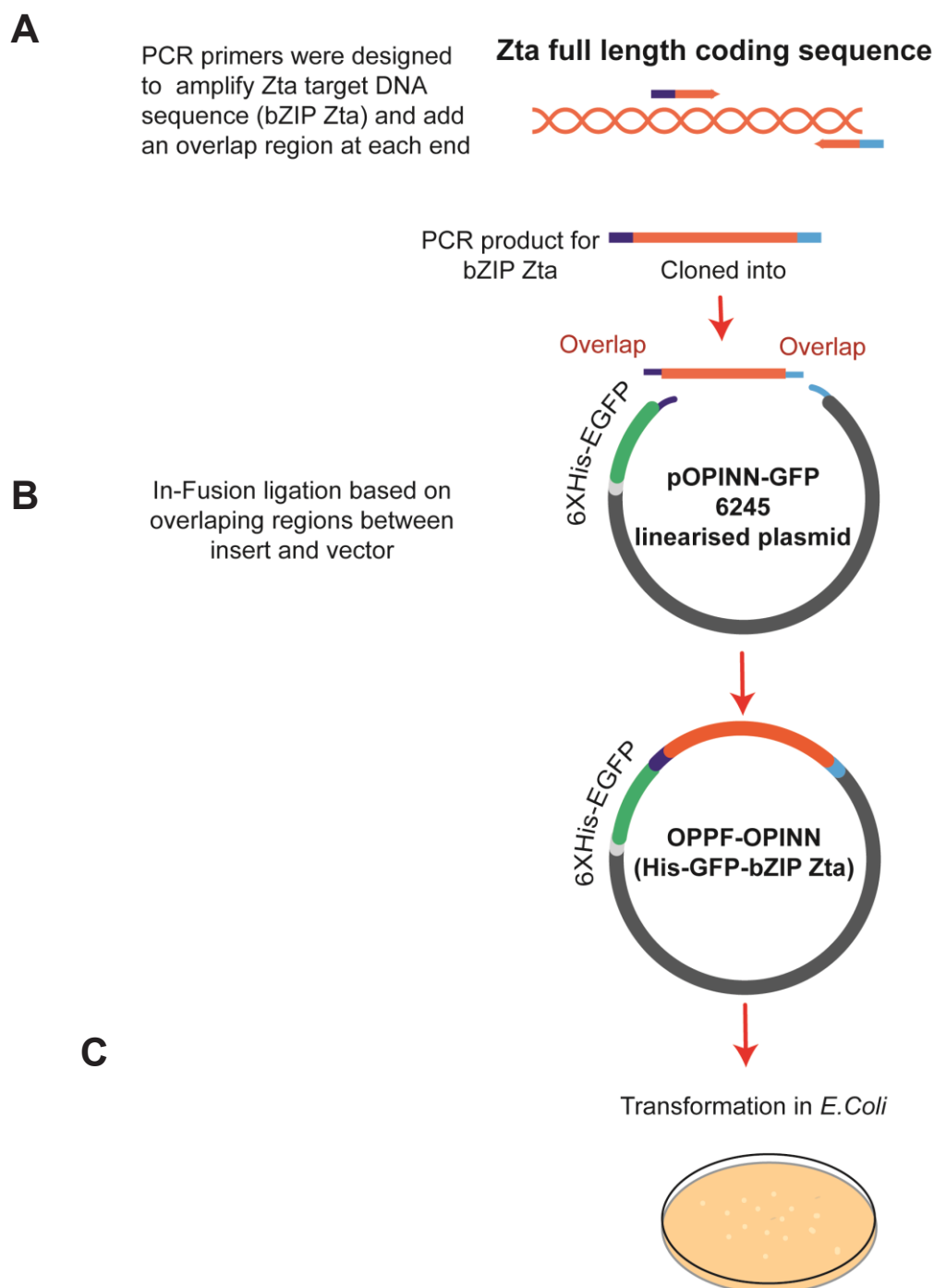
To have a suitable insert to be used in the In-Fusion ligation reaction, PCR primers were designed to add overlapping regions with the expression plasmid. **Figure 5-13** represents a schematic diagram of the cloning process. The PCR product is illustrated along with the overlapping regions (**Figure 5-13 (A)**). The destination plasmid pOPINN-GFP along with the location and orientation of the tag sequence (His-GFP) are shown in **Figure 5-13 (B)**. Subsequently, the

ligated DNA was transformed into competent cells as shown in **Figure 5-13 (C)**. A more detailed description of the In-Fusion ligation is given in **Chapter 2**. The generated plasmid (referred to as His-GFP-bZIP Zta plasmid) was tested for its expression of His-GFP-bZIP Zta recombinant protein in *E. coli*, as previously mentioned.

**Figure 5-14 (A)** shows the DNA sequence for bZIP Zta (168-245 aa) PCR product along with the PCR primers and the cloning overlapping regions. **Figure 5-14 (B)** represents the protein sequence of the expected recombinant protein sequence His-GFP-bZIP Zta (168-245 aa). The tag sequences and cleavage site along with bZIP Zta are highlighted in different colours. Similarly, **Figure 5-14 (C)** represents a schematic cartoon of the expressed recombinant protein His-GFP-bZIP Zta (168-245 aa).

**Figure 5-15 (A-C)** shows a flow diagram of the major steps of this process. The newly generated His-GFP-bZIP Zta plasmid was used to express and purify the recombinant protein His-GFP-bZIP Zta at a large scale in Rosetta pLysS *E. coli*, as described in **Chapter 2**. We undertook a two-step purification approach utilising an affinity column followed by size exclusion chromatography (gel filtration column). The purification steps were carried out in collaboration with Ms Montse Vega (a Master's degree student in our lab).

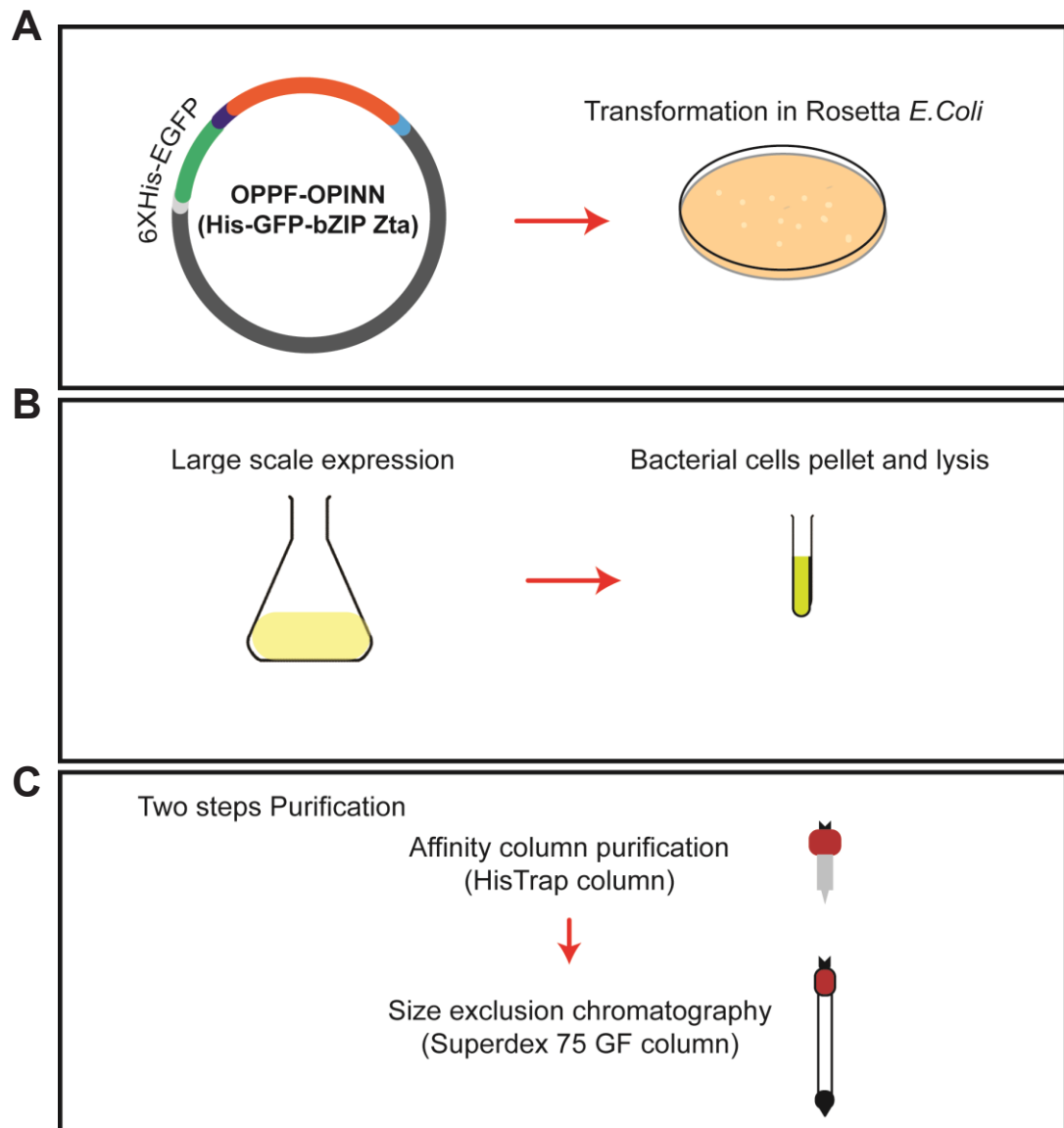
After the protein was eluted from the affinity HisTrap column, a few fractions were pooled together, concentrated and loaded into a size exclusion chromatography column, as described in **Chapter 2**. **Figure 5-16 (A)** shows the UV 280 trace (mAU) of the eluted material after the size exclusion column. The highest-point peak in the graph represents our expressed recombinant protein (His-GFP-bZIP Zta) in its native state (a dimer). A small amount of each collected fraction under the largest peak was separated on a denaturing SDS-PAGE gel stained with SimplyBlue. The band observed immediately above 38 KDa represents the expected size of His-GFP-bZIP Zta **Figure 5-16 (B)**. This enriched band is seen in the fractions under the centre of the main peak, as shown in **Figure 5-16 (A)**. The expression and purification of His-GFP-bZIP Zta resulted in a sufficient amount to be used in BNLF2a ZREs EMSA analysis.



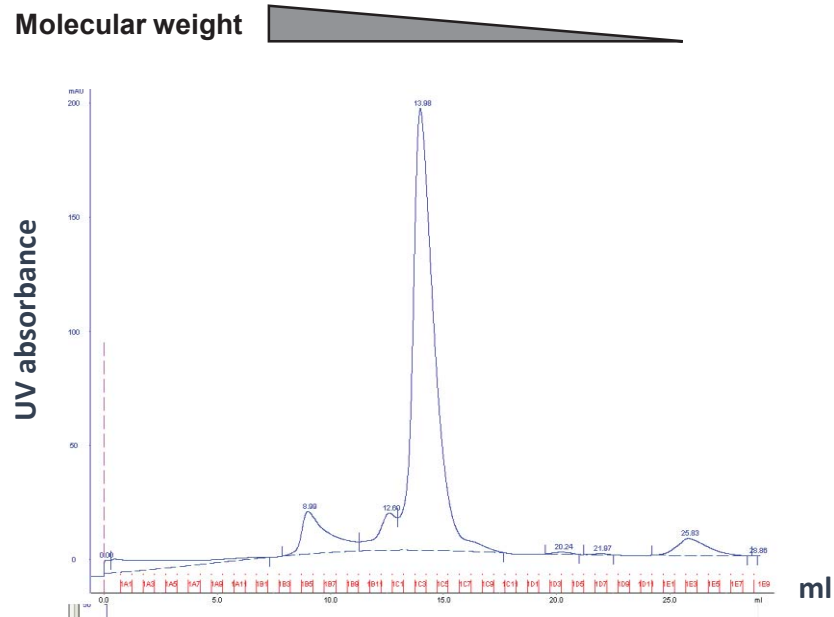
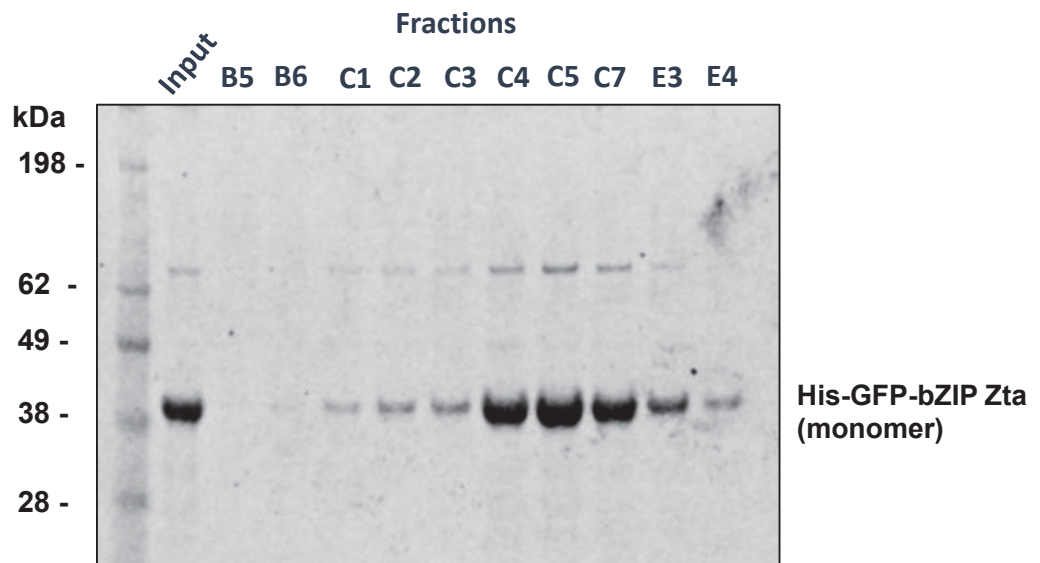
**Figure 5-13 Generating His-GFP-bZIP Zta expression vector.** The bZIP domain of Zta (168-245 aa) tagged with His-GFP at the N terminus was expressed in *E.coli*. The expression vector was generated as a collaboration project with Oxford Protein Production Facility (OPPF). Utilising their high-throughput protein expression technologies, this plasmid as well as other expression plasmids were generated using In-Fusion cloning technology and screened for protein expression. The schematic diagram illustrates the major cloning steps. **(A)** Primers to amplify Zta target DNA sequence were designed with overlap regions (blue colour) between the insert (PCR product) (orange colour) and the expression vector (dark grey) at each end. 6x Histidine-EGFP coding sequence is shown in light grey and green, respectively **(B)** The PCR generated insert and the linearized plasmid were fused together using In-Fusion ligation master mix. **(C)** A transformation in *E.coli* was carried out using the ligated material. All the steps were carried out with the help of Dr Louise Bird at the OPF.

**Figure 5-14 His-GFP-bZIP Zta sequences. (A)** The DNA coding sequence for bZIP Zta (168-245 aa). Primers (forward and reverse) shown in red were designed to amplify Zta target DNA sequence with the addition of overlap regions at each end (dark and light blue colour). The overlapping sequence between the insert (PCR product) and the expression vector (described in previous figure) facilitates the In-Fusion cloning. **(B)** The full protein sequence of the expected expressed protein (His-GFP-bZIP Zta). The N terminus 6 X histidine and GFP tags sequences are shown in dark grey and green, respectively. PreScission protease cleavage site (3C) sequence is underlined in black. Zta bZIP domain sequence (168-245 aa) is shown in red. The protein has a theoretical molecular weight of 38.12 KD **(C)** A cartoon highlighting the recombinant protein sequences in similar colours to the sequence above.





**Figure 5-15. Large scale purification of His-GFP-bZIP Zta protein.** (A) His-GFP-bZIP Zta expression vector ( as previously described in Figure 5-13) was transformed into *E.coli* expression strain (Rosetta pLysS) (B) Transformed Rosetta cells were allowed to grow in Autoinduction media for 16 hours. Subsequently, cells were pelleted and resuspended in lysis buffer. (C) First , the protein was purified using His-tag affinity column and eluted using high imidazole concentration. Then, the purified material was passed through size exclusion column to allow the removal of the imidazole and high salt concentration. All the purification steps were carried out in collaboration with Ms Montse Vega (a Master's degree student undertaking her project in our lab).

**A****B**

**Figure 5-16 Purified His-GFP-bZIP Zta.** (A) The size exclusion chromatography UV<sub>280</sub> trace (mAU) of His-GFP-bZIP Zta. (B) A small amount of each collected fraction (under the largest peak in (A)) was separated on denaturing SDS-PAGE gel stained with SimplyBlue. The fractions under the largest peak were pooled together, aliquoted and stored at -80 C for further use in electrophoretic mobility shift assay (EMSA). The large molecular weight proteins elute first in the size exclusion chromatography.

### 5.2.7.2 *BNLF2a* ZREs interaction with His-GFP-bZIP Zta

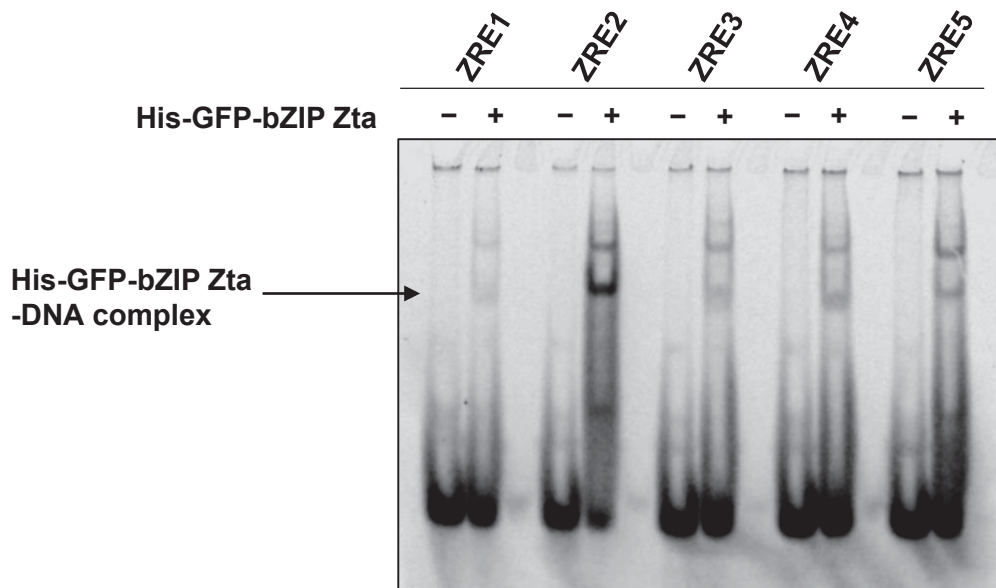
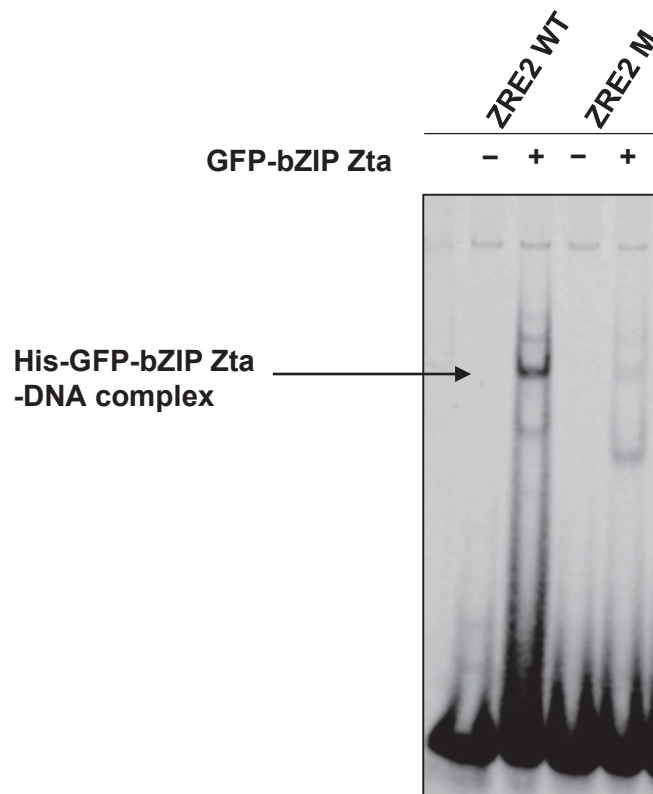
The ability of Zta bZIP domain tagged with GFP to bind *in vitro* to fluorescently labelled probes containing each *BNLF2a* ZRE sequence with a 10 bp flanking on both sides was examined by electrophoretic mobility assay (EMSA). 600 ng of purified protein His-GFP-bZIP Zta was allowed to interact with each probe, then loaded into a DNA retardation gel.

**Figure 5-17 (A)** represents an EMSA gel showing each *BNLF2a* ZRE probe interaction with or without His-GFP-bZIP Zta. A shifted band of the same size was observed for all ZREs, which indicates a His-GFP-bZIP Zta-DNA complex. The binding strength was varied among the different ZREs (ZREs 1-5). ZRE2 showed the strongest affinity to His-GFP-bZIP Zta, as indicated by the increased amount of shifted probe. As previously shown in **Figure 5-3**, ZRE2 is a canonical AP-1 site and found in EBV OriLyt. *BNLF2a* ZREs features and sequences are shown in **Figure 5-3** and **Figure 5-5**.

To confirm if the specificity of the binding between His-GFP-bZIP Zta and ZRE2 was due to the direct interaction with the ZRE core motif, a mutant ZRE2 probe (ZRE2 M) was used (**Figure 5-17 (B)**). In this probe, the core motif of ZRE 2 sequence was replaced to the same mutant sequence used in the luciferase reporter assay analysis (CCCCTTT). A list of all *BNLF2a* ZREs oligonucleotides used to generate the EMSA probes is given in **Table 2-5**.

**Figure 5-17 (B)**, shows the EMSA gel where wildtype ZRE2 probe (ZRE2 WT) was compared with mutant ZRE2 probe. A band shift was observed for ZRE2 WT but not ZRE2 M indicating a specific interaction to the ZRE core motif.

The successful expression and purification of a recombinant Zta protein (His-GFP-bZIP Zta) facilitated the *in vitro* interactions analysis of bZIP Zta with *BNLF2a* ZREs. This analysis revealed a very strong affinity of His-GFP-bZIP Zta to *BNLF2a* ZRE2, while weaker binding was seen for the rest of *BNLF2a* ZREs.

**A****B**

**Figure 5-17 Interaction of Zta with *BNLF2a* ZREs *in vitro*.** (A) The ability of Zta bZIP domain tagged with GFP to bind *in vitro* to fluorescently labelled probes containing each *BNLF2a* ZRE with the 10 bp flanking region on both sides was examined by electrophoretic mobility shift assay (EMSA). 600 ng of purified protein was used with each probe in the presence of Poly dI-dC. (B) The specificity of the binding was investigated by using ZRE2 M probe where the ZRE core motif was changed from TGACACA to CCCCTTT. EMSA retardation gel was imaged using LI-COR imaging system at 800 channel.

### 5.3 Discussion

In agreement with the Zta binding peaks at *BNLF2a* promoter seen in the Zta ChIP-Seq data, five Zta responsive elements (ZREs) were predicted *in silico* upstream of the early lytic gene *BNLF2a* (Ramasubramanyan et al., 2012a). These findings suggested that *BNLF2a* might be a novel target for Zta. Thus, a series of luciferase reporter constructs were generated not only to address Zta activation of *BNLF2a* promoter but also to dissect further which functional ZREs are contributing to *BNLF2a* promoter activation.

The luciferase reporter assay is considered as an efficient, rapid, and highly sensitive technique to investigate regulatory sequences (de Wet et al., 1987; Nguyen et al., 1988). A limitation of this method is the low-throughput outcome, which means that regulatory elements need to be investigated one at a time. Also, this approach does not necessarily reflect the regulation within the chromatin context. However, reporter assays have been employed in global studies by generating thousands of constructs containing variants of regulatory elements downstream of a sequence-specific tagged reporter gene. Upon transient transfection of these constructs into mammalian cells, the specific tagged transcripts for each construct is measured using genome-wide technologies such as RNA-Seq. This approach is known as massively parallel reporter assay (MPRA) (Arnold et al., 2013; Inoue and Ahituv, 2015; Kheradpour et al., 2013; Melnikov et al., 2012; Patwardhan et al., 2012; Sharon et al., 2012).

Unlike most strictly late lytic promoters which are dependent on DNA replication and have a distinct TATA box (Serio et al., 1997; Yuan et al., 2006), *BNLF2a* promoter is similar to other early lytic promoters. These promoters are characterised by the presence of a typical TATA box motif within 30 bp from the transcription start site, mostly regulated by Zta and Rta (Chevallier-Greco et al., 1986) and have multiple ZREs. Interestingly, many of the promoters known to be responsive to Zta contain multiple ZREs. For example, *BZLF1* (Zp) contains two adjacent ZREs (ZIIIA and ZIIIB) (Flemington and Speck, 1990a), and *BRLF1* promoter contains ZRE1, 2 and 3, ZRE 2-3 have a CpG motif (Bhende et al., 2004; Sinclair et al., 1991). Similarly, *BMRF1* promoter contains three

ZREs (Holley-Guthrie et al., 1990), *BALF2* promoter contains at least two ZREs (Hung and Liu, 1999), and *BHLF1* promoter contains four ZREs (Lieberman et al., 1989). Also, we showed 5 ZREs within *BNLF2a* promoter.

One explanation for the role of multiple adjacent sites for Zta comes from a previous study by (Carey et al., 1992). This study showed that Zta acts in a synergistic way to achieve activation. This was suggested when constructs containing either one ZRE or multi ZREs were compared in a chloramphenicol acetyltransferase (CAT) reporter assay. A construct containing one ZRE gave only trace activity, while a construct with two ZREs showed an 18-fold change when compared to the one ZRE construct. A proportional relation between measured transcription and Zta sites (ZREs) was also reported using *in vitro* transcription assay (Carey et al., 1992). A similar finding was reported in (Sato et al., 1992) where Zta failed to activate collagenase and c-Jun promoters which contained a single consensus AP-1 site in a CAT reporter assay. However, adding the AP-1 site (TGAGTCA) and the ZRE motif (TGTGCAA) upstream of the collagenase promoter resulted in Zta activating the promoter (Sato et al., 1992). Also, the work by Carey et al showed that promoters with more ZREs are activated with less Zta concentration. Therefore, more ZRE sites will require less amount of Zta to induce activation.

In *BNLF2a* promoter luciferase assay, I was able to show Zta activation in the presence of all intact ZREs. When all the ZREs were mutated only 15% of the wildtype promoter activity was detected. This observation was consistent in EBV-negative Burkitt's lymphoma B cells (DG75) and cervical carcinoma cells (HeLa cells), as shown in **Section 5.2.4 (Figure 5-8)**. Strikingly, this activity was dependent on the proximal ZREs while the more distal ZREs (ZRE 4-5) had a marginal effect on the promoter (**Figure 5-9**).

Constructs designed to further dissect the promoter, leaving one of the proximal ZREs resulted in much reduced activity, although this reduction was slightly more prominent when ZRE2 was left (**Figure 5-11**). It is worth noting that ZRE1 has a CpG motif, but the methylation effect on Zta transcription activation or binding was not investigated in this project. Zta is known to strongly and

preferentially bind methylated ZREs present in the viral genes *BRLF1*, *BMRF1* and cellular gene *EGR1* promoters (Bhende et al., 2004, 2005; Flower et al., 2011; Heather et al., 2009; Karlsson et al., 2008a; Sinclair et al., 1991).

Leaving out two ZREs in each construct of the *BNLF2a* constructs *BNLF2a* 1&2, *BNLF2a* 1&3, and *BNLF2a* 2&3 rescued most of the wildtype promoter activity (**Figure 5-12**). There was no difference in promoter activation for the combination of ZRE1-ZRE2 and ZRE2-ZRE3. However, the combination of ZRE1 and ZRE3 with mutant ZRE2 resulted in a marginally lower activity compared to the former two combinations. Interestingly, these findings were consistent between *BNLF2a* promoter constructs (*BNLF2a* 2-5, 1&3-5, and 1-2&4-5) and (*BNLF2a* 1&2, 1&3 and 2&3) in spite of mutating ZRE 4-5 in the latter constructs. This also further confirms the minor role for ZRE4-5 in *BNLF2a* activation. The activity level of each generated *BNLF2a* construct is summarised in **Figure 5-18**.

Our data supports a synergy for Zta by binding at two adjacent ZREs, in agreement with a previous report (Carey et al., 1992). This synergy could be dependent on cooperative DNA binding as Zta binding affinities were different for the different ZRE motifs, as was shown in the EMSA analysis (**Figure 5-17**). In a cooperative DNA binding, the role of one transcription factor (TF) would be to interact with a second TF to facilitate its binding to an adjacent binding site (Janson and Pettersson, 1990).

Cooperative DNA binding could be possible through coactivators such as CBP which could indirectly facilitate the interactions between two Zta molecules at adjacent sites. The histone acetyltransferase protein CBP and Zta are known to co-localise at the immediate-early promoters Zp and Rp (Deng, 2003). However, a direct interaction between Zta molecules bound at adjacent sites is also possible.

On the other hand, Zta synergy could be the result of simultaneous contact of multiple Zta molecules with the basic transcription machinery proteins. Thus, the level of activity is possibly proportional to the number of Zta molecules that

are within an appropriate distance to interact with the transcription machinery factors.

It is not conclusive whether ZRE2 plays a greater role in *BNLF2a* regulation despite the observation that ZRE2 had the highest affinity for Zta *in vitro*. Also, *BNLF2a* promoter constructs with mutant ZRE2 showed lower activity than constructs with mutant ZRE1 or ZRE3. We cannot rule out whether the distance between each two ZREs and/or the distance of each ZRE and the TATA box is a factor in *BNLF2a* regulation. The distance between ZRE1-ZRE3 was around 160 bp while it was 78 bp between ZRE1-ZRE2 and 67 bp between ZRE2-ZRE3.

It is generally believed that the cooperative activation of two factors is most effective when the two factors are on the same face of the DNA helix. Also, the distance between the binding sites affects the formation of DNA-protein-protein DNA complex (Huang et al., 2012).

The distance effect on Zta synergistic transcription activation between two Zta binding sites was addressed by Huang et al. In a series of luciferase constructs containing two identical ZREs (ZIIIB) placed 23 bp from a TATA box and separated by increasing distance (adding 2 bp every time), Zta transcription activation was generally reduced with the increasing distance between the two ZREs. However, in a construct where two Zta molecules were bound on opposite faces of the DNA double helix (ZREs separated by 8 bp) showed higher activity than the construct where Zta molecules were bound on the same DNA helix face (ZREs separated by 4 bp) (Huang et al., 2012).

In addition, the distance between the transcription factors and the TATA box influence their activation ability. However, none of the previous reports (Carey et al., 1992; Huang et al., 2012) addressed changes in the distance between the ZREs and the TATA box.

The role for ZRE 4-5 remains unclear as the mutant of ZRE 4-5 had little impact on the promoter activity in both cells (DG75 and HeLa cells) (**Figure 5-9**).



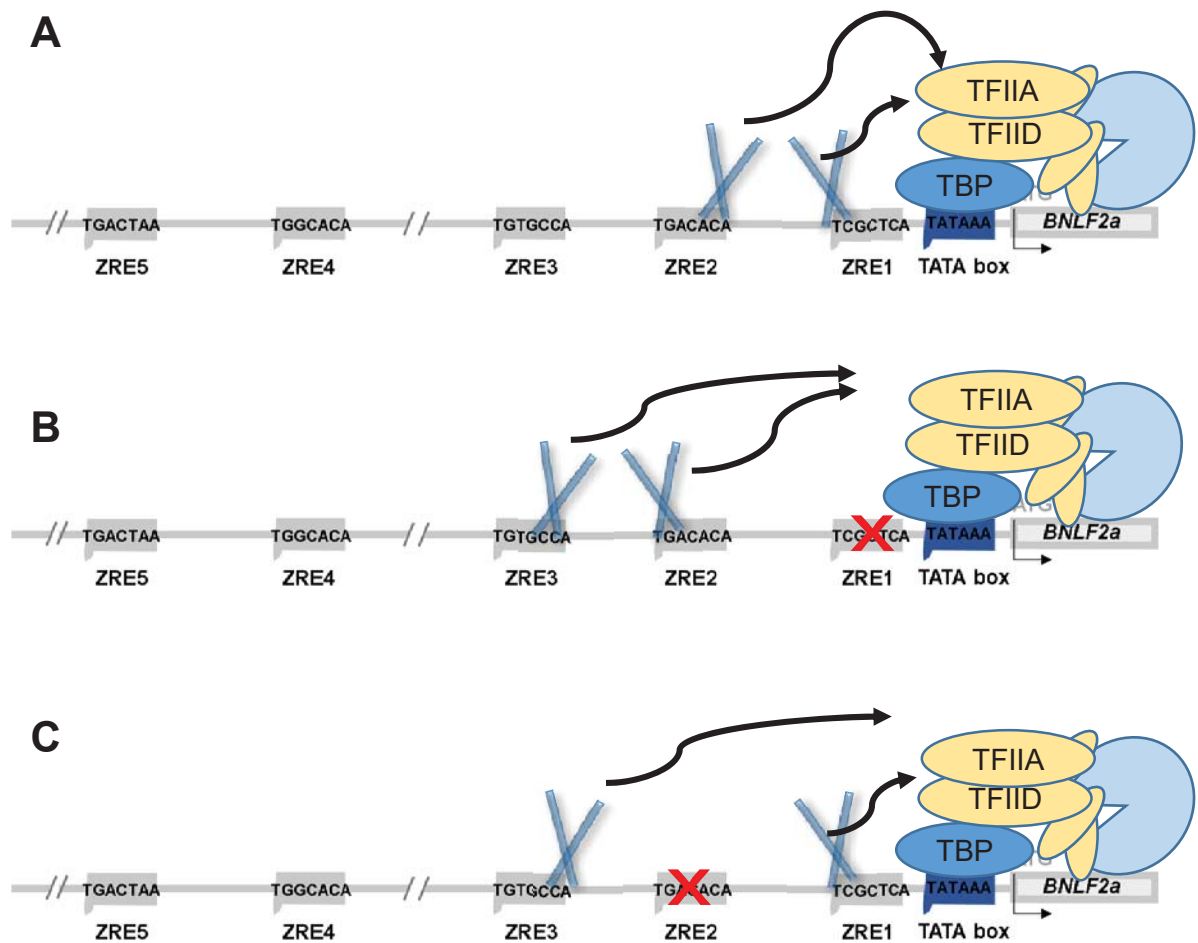
Nevertheless, there was a clear Zta binding peak spanning the region of these two ZREs in the ChIP-Seq data in Akata cells (**Figure 5-1**). In other early promoters such as *BHLF1*, a second ZRE region between -644 to -902 was shown to affect *BHLF1* transcription independently from a more proximal ZREs region (-7 to -155) with some variations in a cell-specific manner (Lieberman et al., 1989).

Based on our luciferase reporter assay and EMSA findings, we propose a speculative model suggesting functional redundancy (fail-safe mechanism) of Zta regulation of *BNLF2a*, explained in **Figure 5-19**. In this model, any two of the proximal ZREs (located within 250 bp upstream of the TSS) can drive promoter activity. However, given the high binding affinity and close distance of ZRE2 to *BNLF2a* TSS, ZRE2 may either be more effective in facilitating the cooperative binding between Zta molecules or the interactions with the TFIID-A complex leading to a more stable preinitiation complex.

*BNLF2a* promoter will be further examined to investigate the role of other regulatory elements, as will be described in the next chapter.

Promoter construct name	Promoter schematic diagram	Contribution to promoter activity in response to Zta
BNLF2a 1-5		+++++
BNLF2a Null		+
BNLF2a 4-5		+ / +++*
BNLF2a 1-3		++++
BNLF2a 2-5		+++
BNLF2a 1&3-5		++
BNLF2a 1-2&4-5		+++++
BNLF2a 1		++
BNLF2a 2		+++
BNLF2a 3		+
BNLF2a 1&2		+++
BNLF2a 1&3		++
BNLF2a 2&3		+++

**Figure 5-18 The ZREs contribution to the promoter activity.** The variant *BNLF2a* promoter luciferase constructs generated to address *BNLF2a* 5 ZREs contribution to the promoter activity are listed along with a schematic representation of each constructs and the activity. \* *BNLF2a* 4-5 showed a slightly more activity in DG75 cells than HeLa cells .



**Figure 5-19 A fail-safe mechanism for *BNLF2a* promoter activation.** Zta transcriptional activity is known to be mediated through Zta response elements as well as interacting with the basal transcription machinery through the TATA binding protein and general transcription factor II D. In *BNLF2a* promoter, the three proximal ZREs are essential for the promoter activity. However, only two are required for activation at the same time where ZRE2 is the key mediator of this activation each time. This is supported by Zta strong binding to ZRE2 *in vitro*. Additionally, the distribution of the three ZREs could be a critical factor as the distance between (ZRE1 and ZRE2) and (ZRE2 and ZRE3) is very similar; 78 and 67 nt, respectively. **Scenario (A):** All ZREs are intact, thus, the promoter is activated through ZRE1 and ZRE2. **Scenario (B):** Failure of ZRE1 leads to Zta being recruited to ZRE3. **Scenario (C):** in case of ZRE2 failure, the promoter is still activated in a less efficient way through ZRE1 and ZRE3. In each case, Zta molecules closest to the TSS interacts with the TFIID-A complex. Zta is represented as blue X. The basal transcription machinery shown as ovals (yellow and blue) RNA Pol II shown as blue circle.

## Chapter 6 Other elements play a role in *BNLF2a* gene regulation

### 6.1 Introduction

As previously discussed in **Chapter 5**, I identified the viral early lytic gene *BNLF2a* as a direct target for Zta. However, to completely understand the mechanisms that govern the regulation of this target and the role of other transcription factors, a comprehensive characterisation of the promoter is required.

*BNLF2a* expression can be detected within 4 hours post-lytic cycle induction, and high expression levels were reported at 24 hours after inducing lytic cycle in Akata BL cells (Tierney et al., 2015). Similarly, *BNLF2a* expression was detected within 4 hours, reached the peak level at 8 hours and remained detectable at 20 hours (Croft et al., 2009) in the Akata derivative cells AKBM (Ressing et al., 2005a).

*BNLF2a* plays a crucial role in evading the host immune response by inhibiting antigen presentation on the cell surface major histocompatibility complex I (MHC I). A comparison between wildtype EBV and delta-*BNLF2a* EBV in LCLs revealed that wildtype EBV-infected cells had a much-reduced recognition by specific CD8<sup>+</sup> T cells to the immediate-early protein BZLF1. This highlighted the critical need for early expression of *BNLF2a* at early stages in EBV lytic cycle (Croft et al., 2009). EBV viral immune evasion and *BNLF2a* role in antigen presentation is previously discussed in **Chapter 1**.

Like most early lytic promoters, *BNLF2a* promoter (also known as EBV ED-L2) contains a canonical TATA-box motif 30 base pairs upstream of the transcription start site (TSS) (Baer et al., 1984; Farrell et al., 1983). This promoter was also shown to be activated by the phorbol ester TPA (Farrell et al., 1983; Jenkins et al., 1997). As previously mentioned in **Section 5.1**, this promoter showed tissue-specific expression in stratifying squamous epithelia (keratinocytes), which are found in the tongue, esophagus, and skin tissues (Nakagawa et al., 1997a; Nakagawa et al., 1997b; Wilson et al., 1990). In addition, *BNLF2a* promoter was employed to express Cyclin D1 in a tissue--

specific manner in a transgenic mice model to investigate oncogenesis in the esophagus (Nakagawa et al., 1997b; Roth et al., 2012). Immunohistochemistry staining for the transgene cyclin D1 showed that expression was restricted to the basal and suprabasal layers of stratified squamous epithelia, but was not detected in simple epithelium layers (Nakagawa et al., 1997b). Likewise, a *BNLF2a* promoter luciferase construct showed constitutive expression of in the human esophageal squamous carcinoma cell line TE-11, but not in other cell lines of simple epithelial origins (Nakagawa et al., 1997a). This has attracted much interest in *BNLF2a* promoter to understand its tissue-specific expression.

A previous mutational analysis of the promoter using luciferase reporter assays showed that a CACCC box-like motif (CACACCT), located between -212 to -105 bp from *BNLF2a* gene start, was critical for basal level expression of *BNLF2a* promoter in TE-11 cells (Nakagawa et al., 1997a). Also, the E-box motif (CACCTG), located between -181 to -175 bp, and a CACACCC motif, located between -132 to -125 bp, were identified to be essential for the phorbol ester TPA induction of the promoter (Jenkins et al., 1997). The CACACCC motif was suggested to be a keratinocyte-specific factor and was shown to form a complex *in vitro* with the kruppel like factor KLF4, also known as gut-enriched kruppel-like factor (GKLF). This was demonstrated by transiently expressing KLF4 which leads to *BNLF2a* promoter activation and by showing direct binding of recombinant KLF4 to the CACACCC motif EMSA experiment (Jenkins et al., 1998).

In this chapter, I focus on analyzing the mechanism of Zta regulation of *BNLF2a*. To achieve this, I examined the promoter for TFs binding sites using an *in silico* approach, and carried out further mutational analysis using luciferase reporter assays. As will be discussed in subsequent sections, several regulatory elements, including a novel motif that we identified, were addressed to shed light on Zta regulation of *BNLF2a* promoter.

## 6.2 Results

### 6.2.1 The conservation of a motif flanking *BNLF2a* ZREs

As previously described in **Chapter 5**, Zta direct regulation of *BNLF2a* through ZREs in the promoter was established.

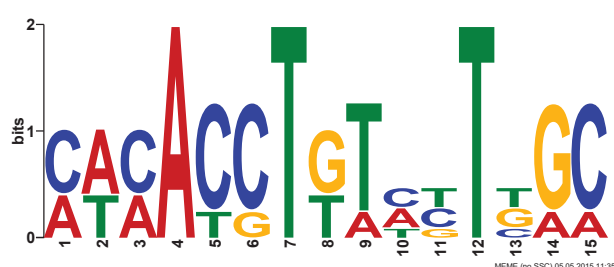
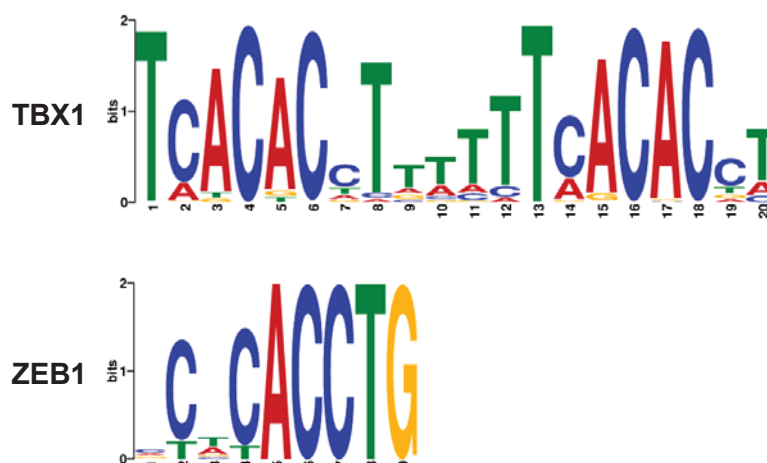
Our specific aim in this chapter was to elucidate a complete model for Zta activation of *BNLF2a* promoter. Therefore, investigating all regulatory elements in the promoter and any possible interacting cellular partners for Zta will be critical. To achieve this, we examined the published literature regarding *BNLF2a* regulation, as described in **Section 6.1**, and conducted *in silico* promoter sequence analysis using the motif discovery tool MEME (Bailey et al., 2009b).

Each *BNLF2a* ZRE sequence along with 14 bp on each side of the ZRE motif was submitted to the MEME motif prediction tool. The submitted sequences are shown in **Figure 6-1 (A)**. Interestingly, this approach predicted a 15 bp ZREs flanking/overlapping sequence which was conserved between all *BNLF2a* ZREs. **Figure 6-1 (B)** shows the position weight matrix of this conserved ZRE flanking motif.

The core bases of the identified motif flanking ZRE1 and ZRE2 overlaps a canonical E-box motif (Atchley and Fitch, 1997). The E-Box flanking ZRE2 is previously identified as one of the essential elements for TPA induction (Jenkins et al., 1997).

This newly identified motif was also submitted to the motif comparison web tool TOMTOM (Bailey et al., 2009b). Interestingly, this identified two significant transcription factors that are known to bind to a similar position weight matrix. These transcription factors are members of T-box binding proteins (TBX) and ZEB1 proteins. **Figure 6-1 (C)** represents the binding site sequence logos of these two transcription factors.

This discovery led to further mutational analysis of *BNLF2a* promoter to understand the role of this motif using luciferase reporter assays as will be discussed in **Section 6.2.2, 6.2.4 and 6.2.5**.

**A****>BNLF2a ZRE1**CTCCACACCTGTCCTCGCTCATCTTTCCACATTCC**>BNLF2a ZRE2**GGAACACCTGTTGTTGACACATTCTTTGCGCATAA**>BNLF2a ZRE3**TTTGCTTTCCATCTTGTGCCAATACACATTTGGAT**>BNLF2a ZRE4**CAAATCACCTTAACTGGCACACACTCCCTTAGCCA**>BNLF2a ZRE5**CTGCTAAGCTACTATGACTAACCTTTCTTTACTTC**B****C**

**Figure 6-1 Position weight matrices (PWMs) of a conserved ZRE flanking motif.** A MEME-predicted ZRE flanking motif is conserved between all *BNLF2a* ZREs. **(A)** *BNLF2a* ZRE sequences along with 14 bp at each ZRE side entered in the MEME motif prediction tool. **(B)** The binding sequence logo revealed by MEME from the entered sequences. **(C)** Submitting the motif shown in (B) to the motif comparison tool TOM-TOM revealed two significant transcription factors match: T-box TBX family and ZEB1. The TOMTOM binding sequence logos for these transcription factors are also shown.

### 6.2.2 Investigating other regulatory elements within *BNLF2a* promoter

To investigate the functional relevance of the various regulatory elements in *BNLF2a* promoter, we undertook the same luciferase reporter assay approach, as described in **Chapter 5**. A further series of *BNLF2a* promoter constructs that did not involve any alteration of Zta response elements (ZREs) were generated using the luciferase plasmid pCpGL-basic.

**Figure 6-2** shows the full DNA sequence (5'-3') of the wildtype *BNLF2a* promoter insert, that was previously described in **Figure 5-5**, this time with each ZRE sequence, the KLF4 site, and the predicted flanking region described in **Section 6.2.1** highlighted. In addition, the canonical TATA motif (30 bp away from the transcription start site) is also highlighted.

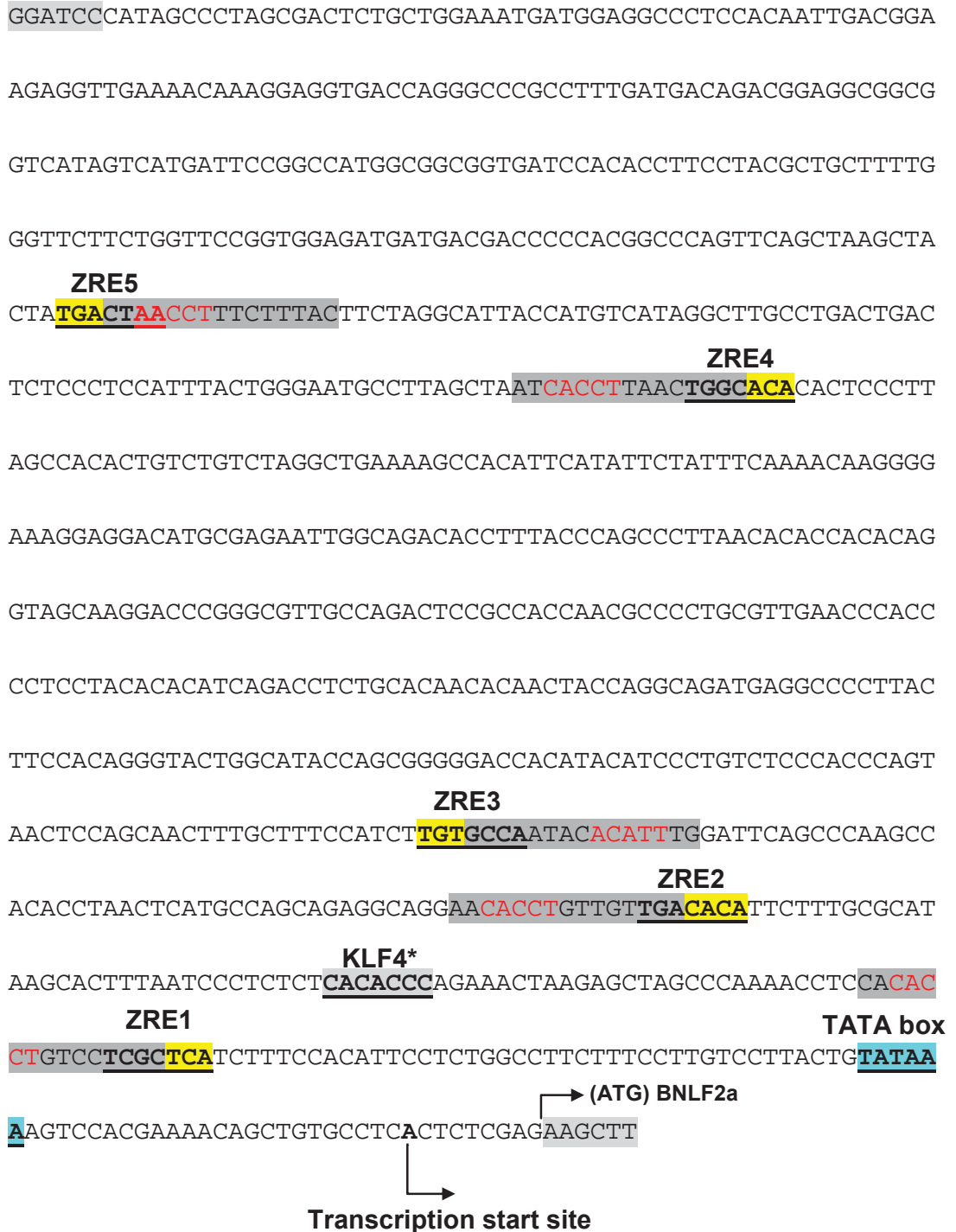
A series of 8 new reporter constructs were generated mutating the highlighted regulatory elements in **Figure 6-2**. These constructs were generated in an identical way to that explained in **Figure 5-4**. In each construct, either one or a different combination of elements were replaced by a random mutant motif, as shown in **Figure 6-3**.

**Figure 6-4** represents a schematic cartoon of the promoter constructs showing the location of each ZRE as well as the other regulatory elements to be examined in this series of experiments.

As in **Chapter 5**, the rationale of each mutant construct will be given in the relevant results section. A list of these mutant constructs including the given name and a brief description of the mutated sites is given in **Figure 6-5**. A small schematic cartoon is also shown for each construct in **Figure 6-5**. This cartoon will be presented again in the relevant results section.

The newly generated reporter constructs were very effective in contributing to our understanding of the possible mechanisms of how Zta activates *BNLF2a*. By utilising luciferase reporter assays, we were able to analyse various cis-acting elements in *BNLF2a* promoter rapidly.





**Figure 6-2 *BNLF2a* promoter and the predicted ZREs flanking motif sequence.** A schematic illustration of the sequence used in *BNLF2a* promoter constructs as described in the previous chapter. A 913 bp sequence with the genome coordinates 167029-167941 based on EBV wildtype strain (NC\_007605.1) was used to analyse *BNLF2a* promoter activity. A MEME motif analysis (described in the previous figure) revealed a conserved motif flanking/ overlapping each ZRE in *BNLF2a* promoter (highlighted in dark grey). The five Zta response elements (ZREs) are highlighted in yellow). In addition, the two motifs described previously in (Jenkins et al.,1998) are highlighted; KLF4 in grey and an E-box motif in red (CACCTG). The E-box is part of the ZRE2 flanking motif. ZRE1 flanking motif is also an identical E-box motif to ZRE2 flank. The canonical TATA box motif (highlighted blue) is located 30 bp upstream the transcription start site. The gene start and transcription start site (TSS) are indicated with arrows. The restriction enzyme sites used in cloning are highlighted in grey.

Sequence Name	Sequence	Mutant sequence
TATA box	<u>T</u> ATA	<u>G</u> ATA
KLF4	CACACCC	ACGTAGA
ZRE1 flanking	CAC <u>ACCT</u> GTCCTCGC	CA <u>GTTTG</u> GTCCTCGC
ZRE2 flanking	AAC <u>ACCT</u> GTTGTTGA	AA <u>GTTTG</u> GTTGTTGA
ZRE3 flanking	GCCAATAC <u>ACATT</u> TG	GCCAATAC <u>CAAAC</u> TG
ZRE4 flanking	AT <u>CACCT</u> TAACTGGC	AT <u>GTTTG</u> TAACTGGC
ZRE5 flanking	CTA <u>ACCT</u> TTCTTTAC	CT <u>GTTTG</u> TTCTTTAC

**Figure 6-3 *BNLF2a* promoter wildtype and mutant elements.** The figure shows the wildtype sequences of the various elements investigated in this chapter. The mutations we introduced in each motif to generate the luciferase reporter assay constructs are also listed (red color).



**Figure 6-4 A schematic diagram of *BNL F2a* promoter showing the ZREs and other elements.** A 913 bp DNA fragment was used to analyse other elements present in *BNL F2a* promoter, as previously described. A schematic illustration of the *BNL F2a* ZREs (light grey boxes) and the other elements are shown. A previously identified KLF4 site (Jenkins et al., 1998) is shown as dark grey box. The TATA box motif is shown as a dark blue box. Green boxes indicate the novel *BNL F2a* ZREs conserved flanking motif described in the previous figure. The numbers of each flank is given according to ZREs numbering as described in the previous chapter (ZREs 1-5). The position of each ZRE (relative to the gene start) is also shown.

Promoter construct Name	Description	Schematic diagram
BNLF2a 1-5	Wildtype sequence	
BNLF2a Null	1-5 ZREs mutant	
BNLF2a TATA M	TATAAA motif replaced with GATAAA	
BNLF2a KLF4 M	KLF4 site (described in (Jenkins et al., 1998) mutant	
BNLF2a 1FL M	ZRE1 flanking motif mutant	
BNLF2a 2FL M	ZRE2 flanking motif mutant	
BNLF2a 3FL M	ZRE3 flanking motif mutant	
BNLF2a 4FL M	ZRE4 flanking motif mutant	

Promoter construct Name	Description	Schematic diagram
BNLF2a 5FL M	ZRE5 flanking motif mutant	
BNLF2a 1-5 FL M	ZREs 1-5 flanking motif mutant	

**Figure 6-5 *BNLF2a* luciferase promoter constructs.** A previously described element (Jenkins et al.,1998), A TATA box variant and the MEME predicted elements were investigated using luciferase reporter assays. The construct given name, mutation site and a schematic diagram of each constructs generated are listed as a table.

### 6.2.3 A single point mutation in the TATA box enhances Zta activation of *BNLF2a*

As a first step towards understanding the various means of Zta transcriptional activation, we examined the effect of changing the TATA motif to GATA motif on Zta activation of *BNLF2a* promoter. This was of great interest as TFIID, in particular, TBP has a weaker affinity to the non-canonical TATA box (GATA). This results in a reduced preinitiation complex recruitment to the core promoter. However, as described in **Section 6.1**, Zta is known to stabilise the general transcription factor binding to GATA motif through direct interaction with TFIID-TFIIA complex, which compensates for the inefficient recruitment of the preinitiation complex (Lieberman and Berk, 1991; Lieberman et al., 1997). Furthermore, a GATA motif is present in one of the highly responsive promoters for Zta, which is the promoter of the early lytic gene *BHLF1* (Lieberman et al., 1989). Thus, we investigated whether or not the non-canonical TATA box (GATA) might provide EBV with a functional transcription advantage over cellular promoters.

A mutant reporter construct (BNLF2aTATA M) was generated containing a GATA motif instead of TATA motif using the luciferase reporter assay plasmid (pCpGL-basic), as described in **Chapter 5**.

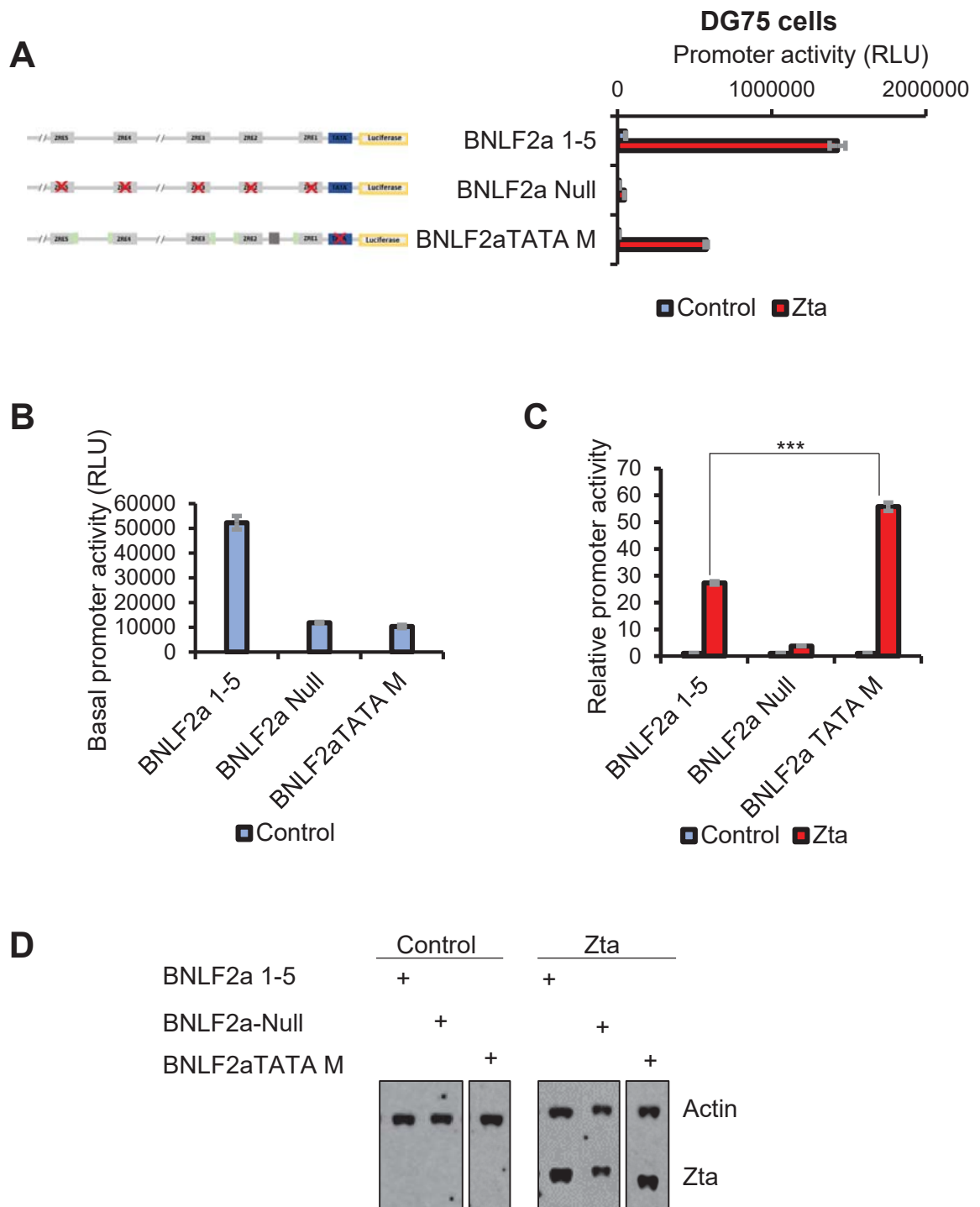
The activity of this construct was compared to the wildtype BNLF2a construct BNLF2a 1-5 and the mutant ZREs construct BNLF2a Null. Each luciferase construct was co-transfected with either the control vector pcDNA3 or His-Zta expression vector into EBV-negative Burkitt's lymphoma cell line (DG75 cells) (**Figure 6-6**) and HeLa cells (**Figure 6-7**).

**Figure 6-6 and 6-7 (A)** represent a comparison between promoter activity for the constructs BNLF2a 1-5, BNLF2a Null and BNLF2aTATA M presented in luciferase relative light units (RLU) for both cell lines. As shown in the previous chapter, Zta activates the promoter in the presence of intact ZREs; however, this activity disappears largely for BNLF2a Null construct, which contains mutant ZREs. The promoter in BNLF2aTATA M construct was significantly activated in the presence of Zta, despite the decrease in the basal promoter

activity observed in the absence of Zta. The basal promoter activity was depicted solely in **Figure 6-6 and 6-7 (B)**. The relative activity of the promoter for BNLf2aTATA M construct, represented by the promoter activity in the presence of Zta over the activity in absence of Zta (control), showed a nearly 2-fold statistically significant increase (P-value <0.001) compared to the relative promoter activity for the wildtype BNLf2a 1-5 construct (**Figure 6-6 and 6-7 (C)**). These findings were consistent between EBV-negative DG75 cells (**Figure 6-6**) and epithelial cells (HeLa cells) (**Figure 6-7**).

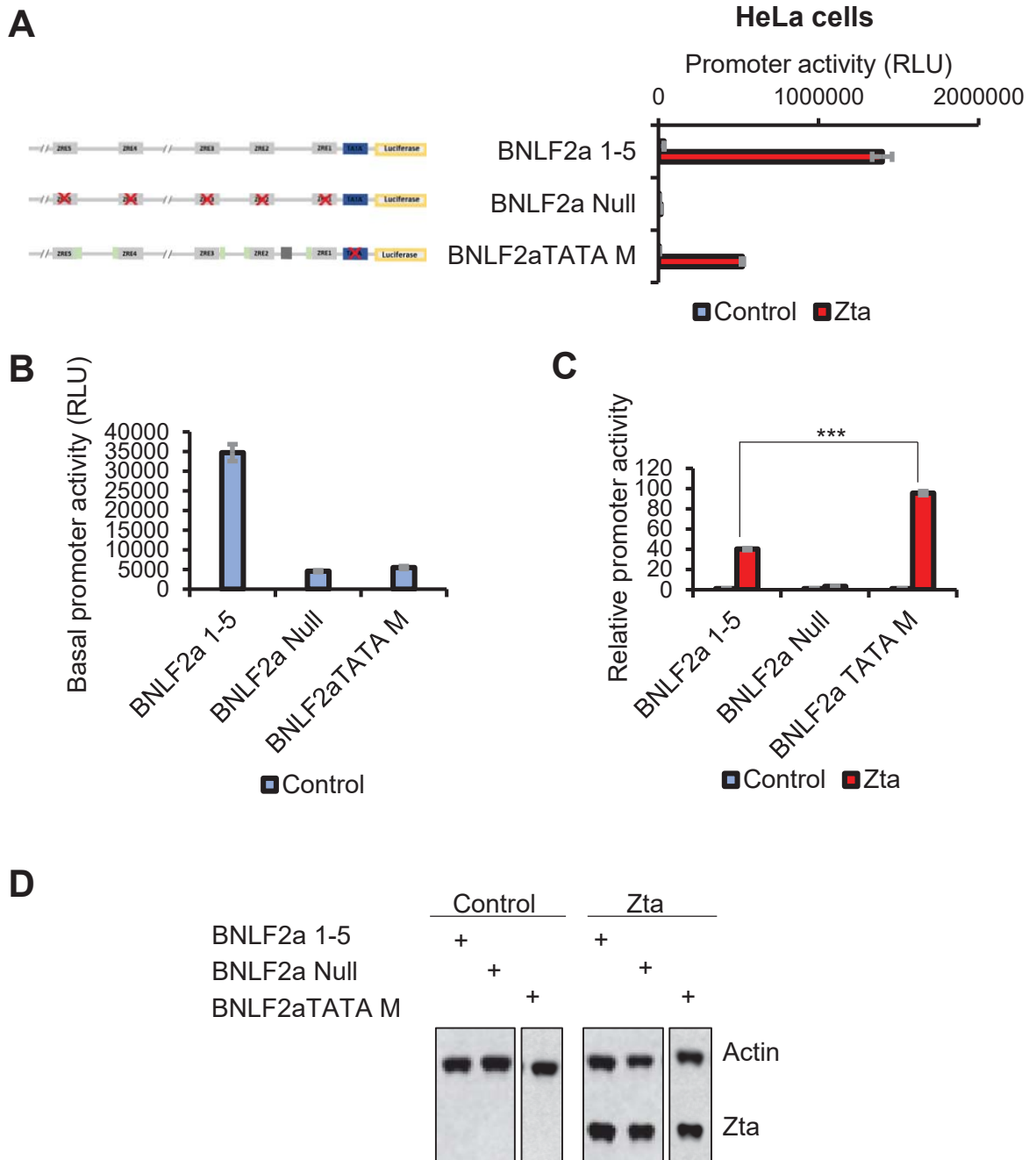
**Figure 6-6 & 6-7 (D)** represents a western blot analysis for actin (loading control) and Zta for each sample. Comparable levels of Zta expression can be seen for all samples while the absence of Zta is confirmed in the control samples. For each sample, the intensity of the actin band was used to normalise the luciferase relative light units value for that sample, as previously described in **Chapter 2**.

The ability of Zta to induce the promoter in the presence of a non-canonical TATA box agrees with the aforementioned published data. However, the more important observation was that the relative promoter activity for BNLf2aTATA M construct was higher than the relative activity of wildtype BNLf2a 1-5. This was in spite of the reduced basal level activity for BNLf2aTATA M.



**Figure 6-6 *BNLF2a* activation in the presence of a non-canonical TATA-box in DG75 cells.** The canonical TATA box motif was changed from TATA to GATA. **(A)** The figure shows the luciferase analysis of the wildtype *BNLF2a* construct (*BNLF2a* 1-5), mutated ZREs construct (*BNLF2a* Null), and TATA-box mutant construct (*BNLF2a* TATA M) co-transfected with either *Zta* expression vector (red bars) or control vector (blue bars) in EBV-negative Burkitt's lymphoma cells (DG75 cells). **(B)** promoter activity in the absence of *Zta* (basal activity) is shown in luciferase relative light units. **(C)** Values are plotted as fold change values over control sample for each construct. Error bars represent the mean of triplicate reads  $\pm$  SD. (\*\*\*) indicates  $p < 0.001$ , compared to *BNLF2a* 1-5 in the presence of *Zta*. **(D)** Western blot shows similar levels of *Zta* transfection in both cell lines.





**Figure 6-7 *BNLF2a* activation in the presence of a non-canonical TATA-box in HeLa cells.** The canonical TATA box motif was changed from TATA to GATA. **(A)** The figure shows the luciferase analysis of the wildtype *BNLF2a* construct (*BNLF2a* 1-5), mutated ZREs (*BNLF2a* Null) construct, and TATA-box mutant construct (*BNLF2a* TATA M) co-transfected with either Zta expression vector (red bars) or control vector (blue bars) in epithelial cells (HeLa cells). **(B)** Promoter activity in the absence of Zta (basal activity) is shown in luciferase relative light units. **(C)** Values are presented as fold change values over control sample for each construct. Error bars represent the mean of triplicate reads  $\pm$  SD. (\*\*\*) indicates  $p < 0.001$ , compared to *BNLF2a* 1-5 in the presence of Zta. **(D)** Western blot shows similar levels of Zta transfection in both cell lines.

#### 6.2.4 The KLF4 site plays a role in *BNLF2a* promoter regulation

As previously described in **Section 6.1**, KLF4 is a zinc transcription factor that was recently associated with Zp and Rp promoters (Murata et al., 2013; Nawandar et al., 2015). A previous report (Jenkins et al., 1998) showed that transient expression of KLF4 is capable of activating *BNLF2a* promoter through a CACACCC motif between ZRE1 and ZRE2, as highlighted in **Figure 6-2**.

To expedite our analysis of *BNLF2a* promoter, we investigated the effect of mutating the CACACCC regulatory element, that was shown to be bound by KLF4 *in vitro* (Jenkins et al., 1998), using a luciferase reporter construct. Thus, the *BNLF2aKLF4 M* construct was generated containing a mutant motif (ACGTAGA) instead of the SP-1 like motif CACACCC.

The activity of *BNLF2aKLF4 M* construct was compared to the wildtype *BNLF2a* 1-5 construct and the mutant ZREs construct *BNLF2a Null*. Each luciferase construct was co-transfected with either the control vector pcDNA3 or His-Zta expression vector into epithelial cells (HeLa cells) (**Figure 6-8**).

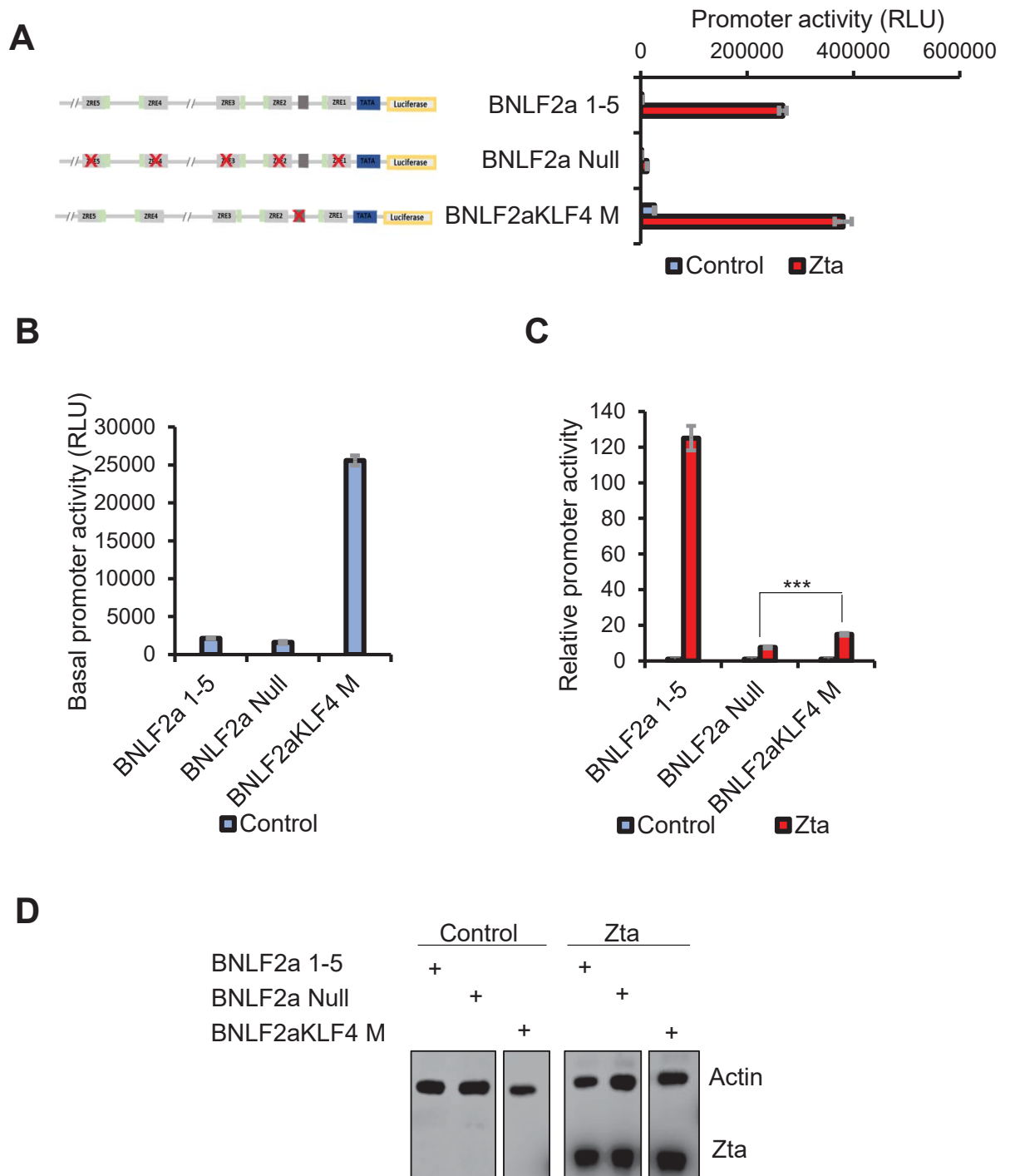
**Figure 6-8 (A)** represents a comparison between promoter activity for the constructs *BNLF2a* 1-5, *BNLF2a Null* and *BNLF2aKLF4 M* presented in luciferase relative light units (RLU) for both cell lines.

As shown in the previous chapter, Zta activates the promoter in the presence of intact ZREs; however, this activity disappears for *BNLF2a Null* construct that contains mutant ZREs. The promoter in *BNLF2aKLF4* construct was activated in the presence of Zta. Remarkably, the promoter basal level has nearly 12-fold increase compared to *BNLF2a* 1-5, in the absence of Zta (**Figure 6-8 (B)**).

The relative activity of the promoter for *BNLF2aKLF4 M* construct, represented by the promoter activity in the presence of Zta over the activity in the absence of Zta (control), showed a substantial decrease in the activation of *BNLF2aKLF4 M* compared to *BNLF2a* 1-5. Nevertheless, this activation was slightly higher than *BNLF2a Null* (P-value <0.001) (**Figure 6-8 (C)**).

**Figure 6-8 (D)** represents the western blot analysis for actin (loading control) and Zta for each sample, as previously described.

By mutating the KLF4 site the basal activity level of the BNLF2a promoter was highly elevated, which suggests a possible role for a repressor factor. These findings are interesting and could suggest a possible diminishing in the ability of Zta to induce the promoter in the presence of mutant KLF4 site, as shown in **Figure 6-8 (C)**.



**Figure 6-8 KLF4 binding site has an effect on basal promoter activity.** A previously identified binding site for KLF4 (Jenkins et al., 1998) is mutated to investigate its effect on *BNLF2a* promoter activation. The motif was changed from CACACCC to **ACGTAGA**. **(A)** The figure shows luciferase analysis of *BNLF2a* wildtype construct (BNLF2a 1-5), mutated ZREs (BNLF2a Null) construct, and KLF4 site mutant construct (BNLF2aKLF4 M) co-transfected with either Zta expression vector (red bars) or control vector (blue bars) in epithelial cells (HeLa cells). **(B)** Promoter activity in the absence of Zta (basal activity) is shown in luciferase relative light units. **(C)** Values are presented as fold change values over control sample for each construct. Error bars represent the mean of triplicate reads  $\pm$  SD. (\*\*\*) indicates  $p < 0.001$ , compared to BNLf2a 1-5 in the presence of Zta. **(D)** Western blot shows similar levels of Zta transfection in both cell lines.

### 6.2.5 Mutational analysis of the motif flanking ZREs revealed a repressor binding site

This section shows the investigation of the *in silico* predicted motifs flanking BNLF2a ZREs, as described in **Section 6.2.1**. We investigated this highly conserved motif which is either adjacent to or overlaps with the ZREs within the BNLF2a promoter. Further analysis using TOMTOM transcription factor comparison web tool identified the transcription factors TBX family and ZEB1 as potential binding factors to this flanking motif.

To analyse this motif flanking the ZREs a similar way to the previous luciferase reporter analysis, 5 new constructs (BNLF2a 1FL M - BNLF2a 5FL M) were generated containing mutations in each motif. Additionally, a construct containing a combination of all mutant motifs was generated (BNLF2a 1-5FL M).

A list of all the generated constructs in this chapter was given previously in **Figure 6-5**. In addition, the mutant motif sequences for each construct are also shown in **Figure 6-4**.

**Figure 6-9** and **Figure 6-10** shows the activity of these constructs compared to the wildtype BNLF2a 1-5 and BNLF2a Null constructs. Each luciferase construct was co-transfected with either a control vector (pcDNA3) or His-Zta expression vector in HeLa cells.

**Figure 6-9 (A)** represents the promoter activity for all mutant flanking motif constructs compared with the BNLF2a 1-5, BNLF2a Null in normalised luciferase light units (RLU).

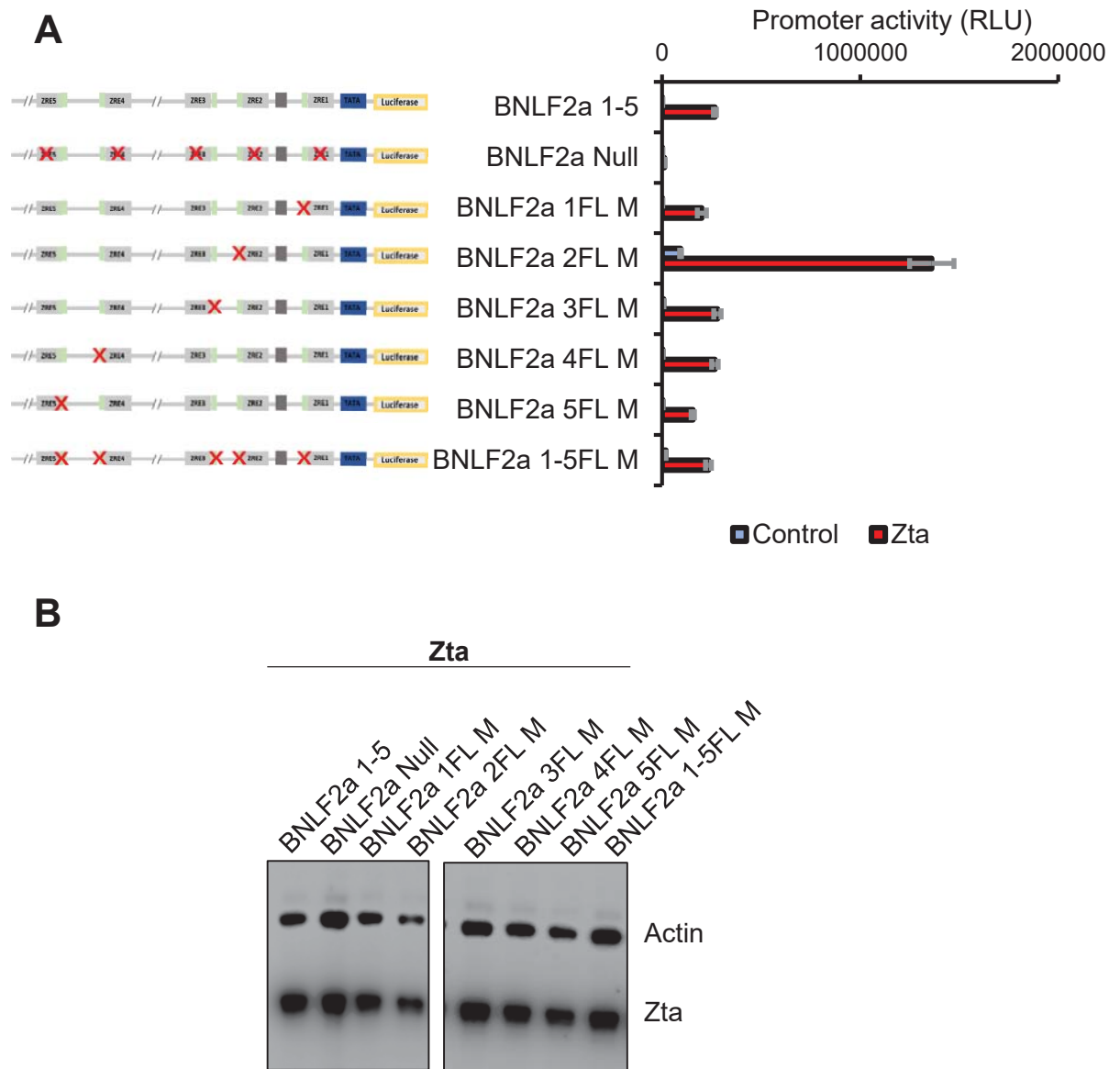
For the motif constructs, there was no single construct that completely impaired Zta activation. Unlike the rest of the mutant motif constructs, BNLF2a 2FL M had nearly five times more activity when compared to BNLF2a 1-5 in the presence of Zta. This motif is only 3 nucleotides upstream of ZRE2 and resembles a canonical E-box motif (Atchley and Fitch, 1997). Interestingly, the basal activity level in this construct was around 44-fold increase than the basal

activity level in BNL2a 1-5 construct. This can be clearly seen in **Figure 6-10 (A)**, where the basal promoter activity was plotted solely. Although, BNL2a 1-5FL M construct which contains a mutant sequence for all flanking motifs, it did not show similar basal level activity to BNL2a 2FL M construct.

**Figure 6-10 (C)** shows the relative activity of the promoter for all ZREs flanking motif constructs, represented by the promoter activity in the presence of Zta over the activity in the absence of Zta (control). Noticeably, this showed a reduced activation in BNL2a 2FL M and BNL2a 1-5FL M when compared to BNL2a 1-5 construct. However, the activation levels for these two constructs were slightly higher than BNL2a Null (P-value < 0.001).

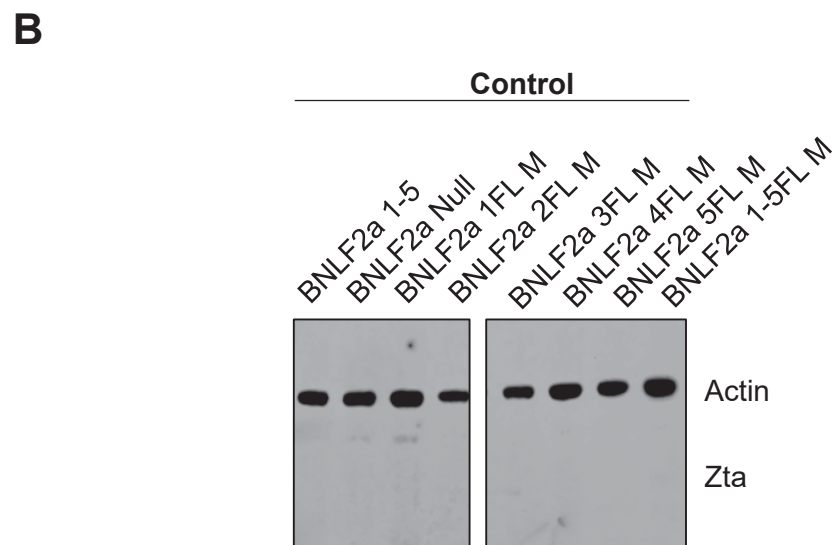
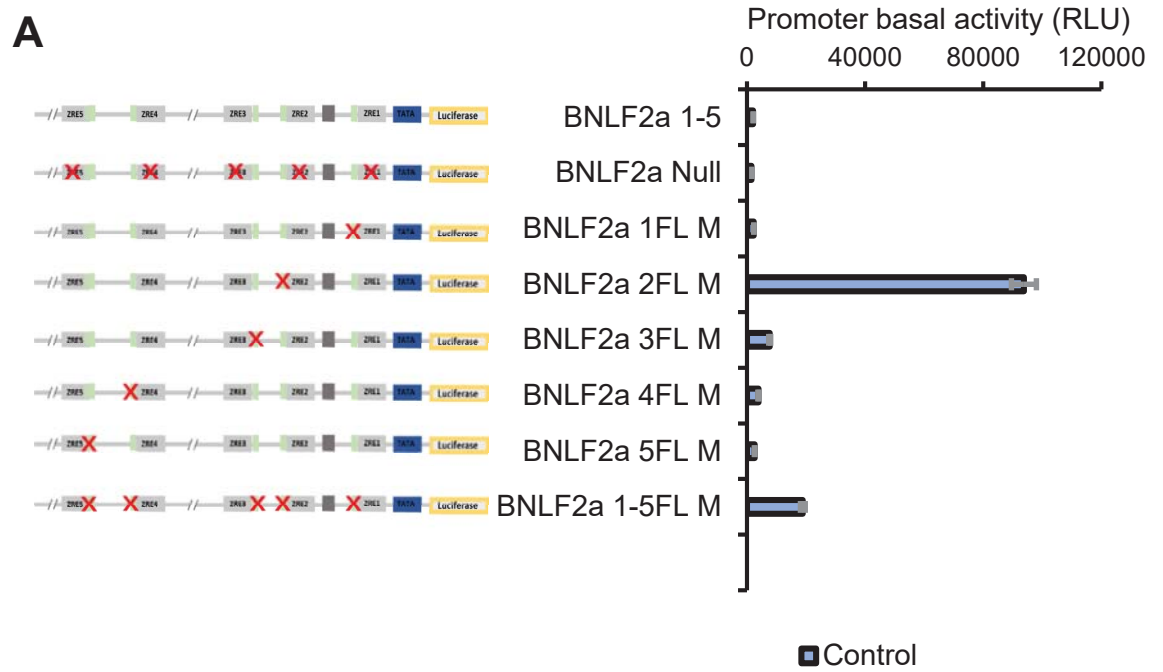
**Figure 6-9 and Figure 6-10 (B)** show a western blot analysis blotted for a loading control (actin) and Zta for each sample is shown as previously described.

Our luciferase reporter analysis of the ZREs flanking motif has revealed an important site next to ZRE2. By mutating this site, a marginal increase in *BNL2a* promoter basal level was reported which strongly suggests a possible role for a cellular repressor. These findings along with the prominent role observed for ZRE2, as discussed in **chapter 5**, will require further investigation to elucidating Zta activation mechanism of *BNL2a*.

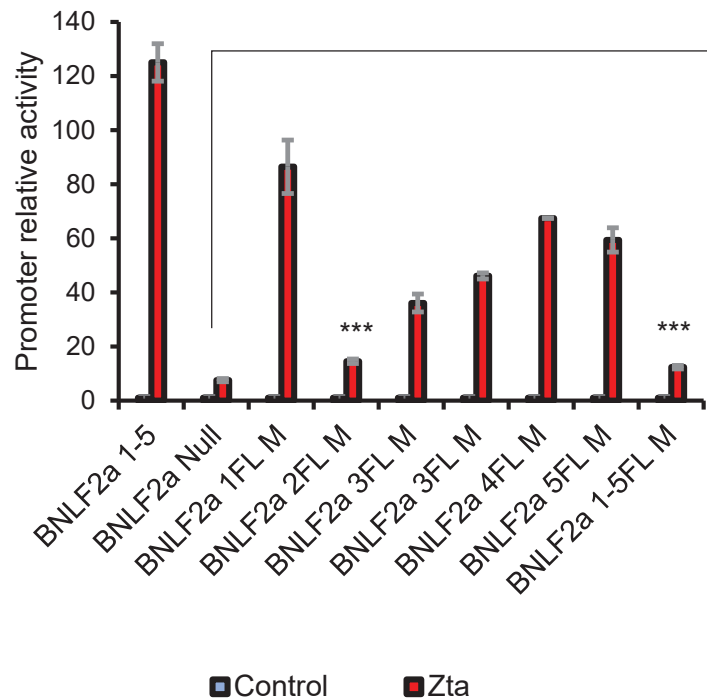


**Figure 6-9 ZREs flanking motif effects on Zta activation of *BNL2a* promoter.**

(A) The figure shows luciferase analysis of BNL2a wildtype construct (BNL2a 1-5), mutated ZREs (BNL2a Null) construct as well as flank mutants constructs. *BNL2a* promoter luciferase constructs were co-transfected with either Zta expression vector (red bars) or control vector (blue bars) in epithelial cells (HeLa cells). ZRE2 flank mutant (E-box motif) shows high level of activation compared to BNL2a 1-5 (wildtype). (B) Western blot shows similar levels of Zta transfection for Zta samples. Values are normalised to actin levels and represented as luciferase relative light units. Error bars represent the mean of triplicate reads ± SD.





**C**

**Figure 6-10 ZRE2 flanking motif has a profound effect on promoter basal activity.** The figure shows the basal promoter activity in HeLa cells of BNLf2a wildtype construct (BNLf2a 1-5), mutated ZREs (BNLf2a Null) construct as well as constructs where the ZREs flank are mutated. Mutation of an E-box flanking ZRE2 increased the promoter basal level (**B**). Western blot shows the loading control actin for control samples. Control values are normalised to actin levels and represented as luciferase relative light units as seen in the previous page. (**C**) Values are presented as fold change values over control sample for each construct. Error bars represent the mean of triplicate reads  $\pm$  SD. . (\*\*\*) indicates  $p < 0.001$ , compared to BNLf2a Null in the presence of Zta.

### 6.2.6 Zta binding to ZRE2 is not affected by mutation in the flanking motif

As described in **Section 5.2.7**, ZRE2 had the strongest binding affinity to Zta bZIP domain *in vitro*. Given the close proximity of the newly identified flanking motif to ZRE2 core motif, the question of whether Zta binding is affected by the flanking motif or not was addressed.

Therefore, Zta binding to ZRE2 was investigated *in vitro* with an EMSA probe containing the wildtype or mutated flanking motif (E-box motif). Although, the EMSA experiment described here was carried out in largely a similar way to the EMSA experiment described in **Section 5.2.7**, different probes for ZRE2 were used (**Figure 6-11 (A)**). This was because in the previous probes, the E-box motif was at the end of the probe used in the previous experiment, which could give false negative results. Thus, new probes were designed to ensure that both ZRE2 and the flanking E-box motif is within acceptable distance from the ends of the probe. Accordingly, fluorescently labelled probes were designed and ordered as described in **Chapter 2**; these new probes are referred to as BNLF2a ZRE2 E-BOX probes. The sequences for all EMSA probes are listed in **Table 2-5**.

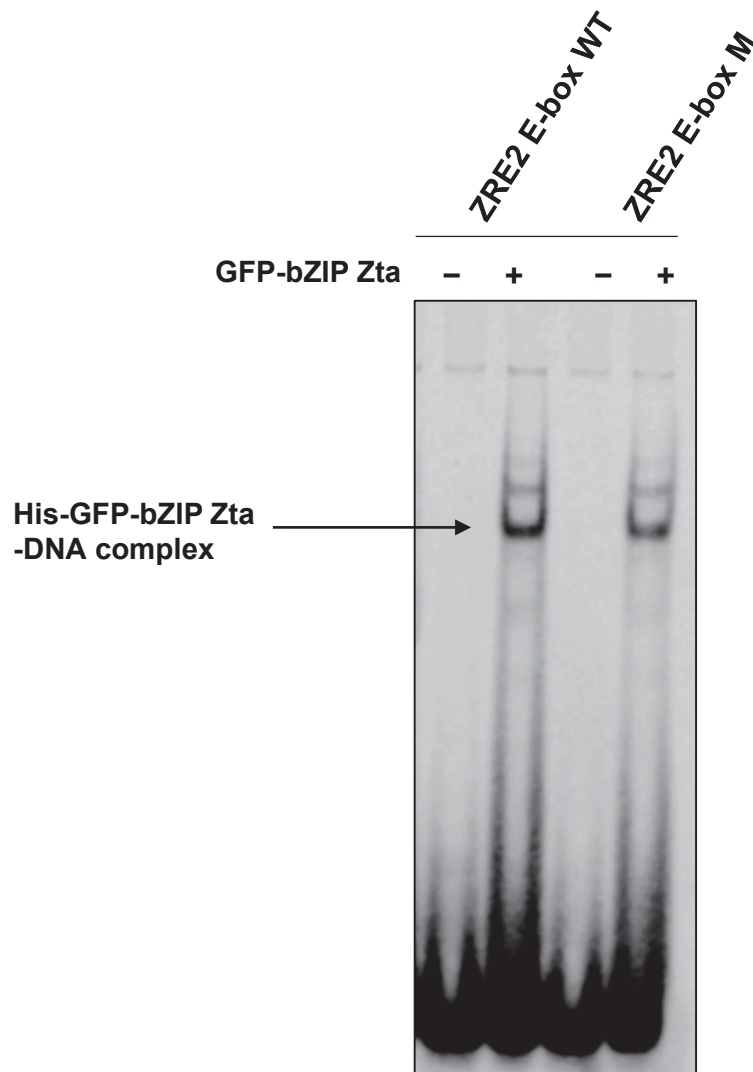
600 ng of a recombinant Zta protein resembling His-GFP tagged basic leucine zipper (bZIP) domain of Zta (His-GFP-bZIP Zta) was used in our EMSA analysis as previously described in **Section 5.2.7**. **Figure 6-11 (A)** shows the 5'-3' strand of the probe used in this experiment with the ZRE and E-box flanking motif highlighted. **Figure 6-11 (B)** represents an EMSA gel showing the BNLF2a ZRE2 E-box and BNLF2a ZRE2 E-box M interaction with or without His-GFP-bZIP Zta.

A band of the same size representing the shifted DNA-protein complex was observed for both probes. The very small difference observed in the shifted band for ZRE2 E-BOX M was not taken into account as the ratio of the shifted band intensity over the free probe band was similar for both probes.

In conclusion, a purified His-GFP-bZIP Zta was shown to bind to ZRE2 independently of the adjacent E-box motif (ZRE2 flanking motif).

**A**

BNLF2a ZRE2 E-box <b>WT</b>	AGGAACACCTGTTGT <b>TGACACA</b> TTCTT
BNLF2a ZRE2 E-box <b>M</b>	AGGAAGTTTGTTGT <b>TGACACA</b> TTCTT

**B**

**Figure 6-11 Interaction of Zta with the mutant motif flanking ZRE2.** The ability of Zta bZIP domain tagged with GFP to bind *in vitro* to fluorescently labelled probes containing a mutant of the E-box motif adjacent to BNLF2a ZRE2 ( 3 bp away) was examined by electrophoretic mobility assay (EMSA). **(A)** The 5'-3' sequence of ZRE2 flanking motif EMSA probes is shown. The wildtype and mutant motifs of the E-box are coloured in red. ZRE2 motif is given in blue. **(B)** 600 ng of purified protein was used with each probe in the presence of Poly dl-dC. The E-box (ZRE2 flanking core motif) was changed from CACCTG to GTTTGG. EMSA retardation gel was imaged using LI-COR imaging system at 800 channel.

### **6.2.7 Specific complex binds to ZRE2 flanking sequence (E-box) in HeLa cells nuclear extract**

Given the effect of mutating the E-box motif flanking ZRE2, as described in **Section 6.2.5**, we attempted to identify any possible target protein that may bind to this motif in HeLa cell nuclear extract.

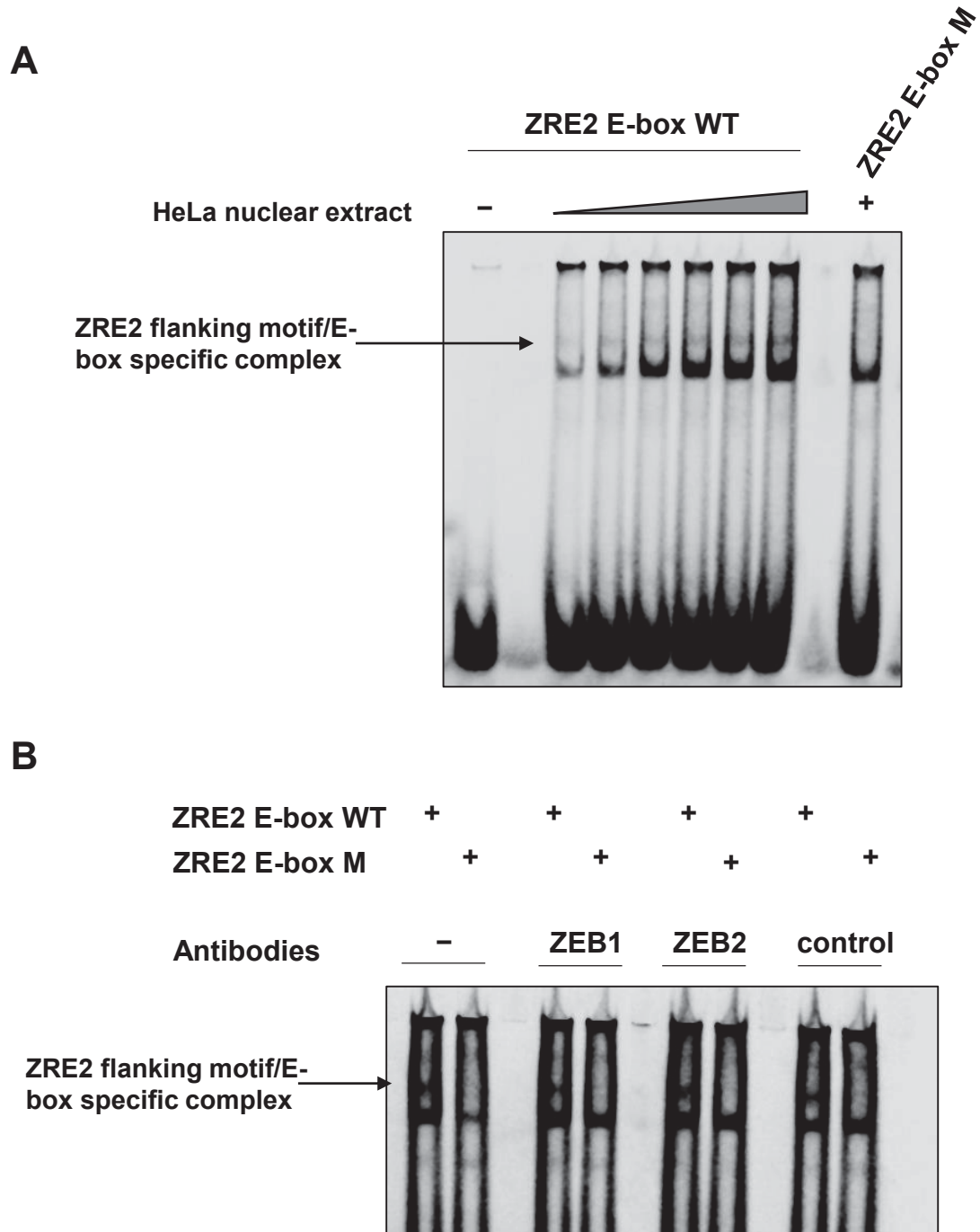
Using EMSA approach, we incubated the BNLF2a ZRE2 E-box and BNLF2a ZRE2 E-box M, as described in the previous section, with increasing amount of nuclear extract (**Figure 6-12 (A)**).

Two shifted bands were observed indicating two different complexes binding to this probe. The upper complex was found to be specific to the flanking motif (E-box site) as this complex disappeared in the probe with mutated E-box (BNLF2a ZRE2 E-box M).

We hypothesised that the two-related zinc finger E-box binding homeobox proteins ZEB1 and ZEB2 could be potential proteins binding to the E-box motif. ZEB1 and ZEB2 are known to repress Zta promoter (Zp) through E-box sites in ZV and ZV' elements within Zp (Ellis et al., 2010). Therefore, we investigated if a supershift band will form with any of the antibodies against ZEB1 or ZEB2 (**Figure 6-12 (B)**).

There was no observation of any supershift band or the disappearance of the shifted band for both ZEB1 and ZEB2 antibodies, as shown in **Figure 6-12 (B)**.

Thus, the specific complex binding to the motif flanking ZRE2 in HeLa nuclear extract remains to be identified.



**Figure 6-12 EMSA using HeLa nuclear extract.** HeLa nuclear extract was incubated with the fluorescently labelled wildtype and mutant ZRE2 flanking motif (E-box) probes to examine for any DNA-protein complex presence using EMSA **(A)** Increasing amount by 5  $\mu$ g of HeLa nuclear extract was used. 30  $\mu$ g was used with ZRE2 E-box M probe. **(B)** 2  $\mu$ g of each different antibody was used to test for potential binding targets (ZEB1 and ZEB2). Zta antibody (SCZ) was used as a control antibody. Super shift or disappearance of the shifted band was not observed. An arrow points to a specific complex to the flanking motif. EMSA retardation gel was imaged using LI-COR imaging system at 800 channel.

### 6.3 Discussion

In the previous chapter, I showed Zta regulation of *BNLF2a* was through the direct binding to ZREs within the promoter using luciferase reporter assays. Here, I extend the analysis of the promoter to identify new regulatory elements and address the role of various other motifs in the context of Zta regulation.

Initially, we examined the TATA motif, as mentioned in **Section 6.2.3**. a TATA variant motif (GATA) is present in *BHLF1* promoter which is the strongest known EBV promoter in terms of Zta activation (Lieberman et al., 1989). Also, it is one of the most abundant transcripts during lytic (Metzenberg, 1989). Zta binding partner Ku80 has been reported to enhance Zta activation of this promoter (Chen et al., 2011). *BHLF1* transcript is essential to initiating EBV lytic DNA replication (Rennekamp and Lieberman, 2011). However, the benefit of having a non-canonical TATA box motif (GATA) in *BHLF1* regulation remains unclear.

In the cellular gene erythropoietin, the promoter contains a highly conserved non-canonical TATA-box (GATA). This motif plays a role in selectively expressing this gene during hypoxia, while the promoter remains repressed in normoxia (Tsuchiya et al., 1997). A similar finding was reported in the chicken beta-globin promoter where a GATA motif instead of a TATA motif (referred to as a specialised TATA element), was found to restrict the expression of this gene in erythroid cells. A mechanism was proposed in which an erythroid-specific protein (GATA-1) facilitated the transcription of this promoter with the aid of another GATA-1 binding site at 3' enhancer element (Fong and Emerson, 1992; Penner and Davie, 1994).

Although, the TATA-binding protein (TBP) recruitment, a rate-limiting step in transcription activation (Xiao et al., 1995), does not form a stable complex with a GATA motif (Lieberman et al., 1997; Tsuchiya et al., 1997), Zta is known to stabilize the TFIID-A complex binding to GATA motif (Lieberman, 1994; Lieberman and Berk, 1991; Lieberman et al., 1997). I used *BNLF2a* promoter as a model system to probe the role of GATA versus TATA on transcriptional activation. To ask if a GATA motif would provide any advantages in terms of Zta

promoter activation, a single point mutation in *BNLF2a* promoter TATA motif was introduced as explained in **Section 6.2.3**. As shown in **Figure 6-6 and 6-7**, when the TATA motif was replaced with a GATA motif, the basal level was reduced compared to *BNLF2a* 1-5 construct, which is expected given the low affinity of TBP to GATA motif. However, in the presence of Zta, the promoter activity was restored confirming the ability of Zta to stabilise the TFIID-A complex, thus initiating transcription. Interestingly, comparing the fold change between *BNLF2a* 1-5 and *BNLF2aTATA M*, the promoter activity of *BNLF2aTATA M* was at least twice the activity in *BNLF2a* 1-5 in both DG75 and HeLa cells. Thus, a GATA motif might provide a more efficient mechanism to restrict and highly express viral genes during EBV lytic cycle.

In **Figure 6-8**, the role of a CACACCC motif (referred to as KLF4 site) located between ZRE1 and ZRE2 was investigated. As described previously in **Section 6.1**, a Kruppel-like factor (KLF4) recombinant protein was shown to bind this motif. Furthermore, transient expression of KLF4 leads to *BNLF2a* promoter activation in 293 cells (Jenkins et al., 1998). KLF4 was also employed in inducing EBV lytic cycle by binding and upregulating *BZLF1* promoter (Zp) and *BRLF1* promoter (Rp) (Murata et al., 2013; Nawandar et al., 2015). KLF4, a zinc finger transcription factor related to Sp proteins, is highly expressed in the epidermal layer of the skin and epithelial cells of the tongue, esophagus and stomach (Garrett-Sinha et al., 1996) and it plays a critical role in epithelial cells proliferation and differentiation (Yori et al., 2010). Knockout mice harboring a deletion for KLF4 fail to develop skin and die shortly after birth (Segre et al., 1999). KLF4, which binds to a GC-rich motif or a CACCC motif, is known to activate various promoters such as the human Keratin 4 promoter through direct binding (Okano et al., 2000). On the other hand, KLF4 was implicated in negatively regulating other promoters such as Cyclin D1 through binding to an SP-1 motif (GGGCGGG). Competition between SP-1 and KLF4 was suggested to be critical for this negative regulation (Shie et al., 2000).

In the luciferase reporter assay data shown in **Figure 6-8**, mutating the KLF4 site resulted in a substantial increase in *BNLF2a* promoter basal activity in HeLa cells. This indicates that this site is possibly bound by a repressor protein.

Although KLF4 has a role in repressing certain promoters, BNLF2a KLF4 site is likely to be occupied by a different factor in HeLa cells. This is based on the fact that transient expression of KLF4 was shown to upregulate the *BNLF2a* promoter in 293 cells.

KLF4 is known to be expressed in HeLa cells (Yang and Zheng, 2012). However, competition with other factors from the SP/KLF proteins family cannot be eliminated as many SP/KLF factors bind to similar motifs. Many KLF proteins are known to have a repressor domain or negatively regulates target promoters by recruiting co-repressors. For example, KLF8 and KLF3 associates with the co-repressor CtBP (Kaczynski et al., 2003; Vliet et al., 2000). Given the fact that Sp proteins have lower affinity to CACCC motif than other KLF proteins (Kaczynski et al., 2003). Such competition between Kruppel like factors has been reported in other cell lines. For example, KLF3 was shown to repress KLF1 target promoters in erythroid cells (Funnell et al., 2012). Even though an attempt to identify an overlapping motif with the KLF4 CACCCC or adjacent repressor sites was not fruitful, such possibility of the presence of an overlapping motif cannot be completely ruled out.

It is worth noting that in the *BZLF1* promoter (Zp), KLF4 was shown to activate the promoter through the SP1 binding sites present in (ZIA, ZIC, and ZID elements) (Murata et al., 2013). On the contrary, KLF4 is reported to activate *BRLF1* promoter (Rp) through two different KLF4 consensus motifs (CACCC and GGGTG) separated by nearly 50 bp. Mutating one of the sites resulted in partial decrease in the promoter activity (Nawandar et al., 2015). This indicates a possible synergistic role for KLF4 in activating these promoters.

Another point to consider in **Figure 6-8**, by examining the promoter fold change activity in **Figure 6-8 (C)**, Zta activation was impaired in the KLF4 M construct compared to the wildtype (BNLF2a 1-5). This suggests a critical role for this element in Zta activation.

In conclusion, CACCC acts as a repressor element in HeLa cells and affects Zta activation largely. However, addressing how this element affects Zta, what



repressor TFs are likely to be involved, and what role does KLF4 play in the promoter activation will require more investigations.

On the other hand, the promoter was subjected to *in silico* motif search through the web tool MEME followed by TOMTOM search as was discussed in **Figure 6-1**. The MEME search tool is useful to compare different DNA sequences/elements (e.g. promoters) that are believed to be regulated in the same way or at least contain the same binding motifs (Bailey et al., 2009b). MEME sequence analysis of BNLF2a ZRE 1-5 revealed a 15 bp conserved motif flanking all BNLF2a ZREs. This flanking motif overlaps a canonical E-box motif in both ZRE1 and ZRE2 flanks. The E-box adjacent to ZRE2 was previously shown to be important for TPA induction of the promoter (Jenkins et al., 1997).

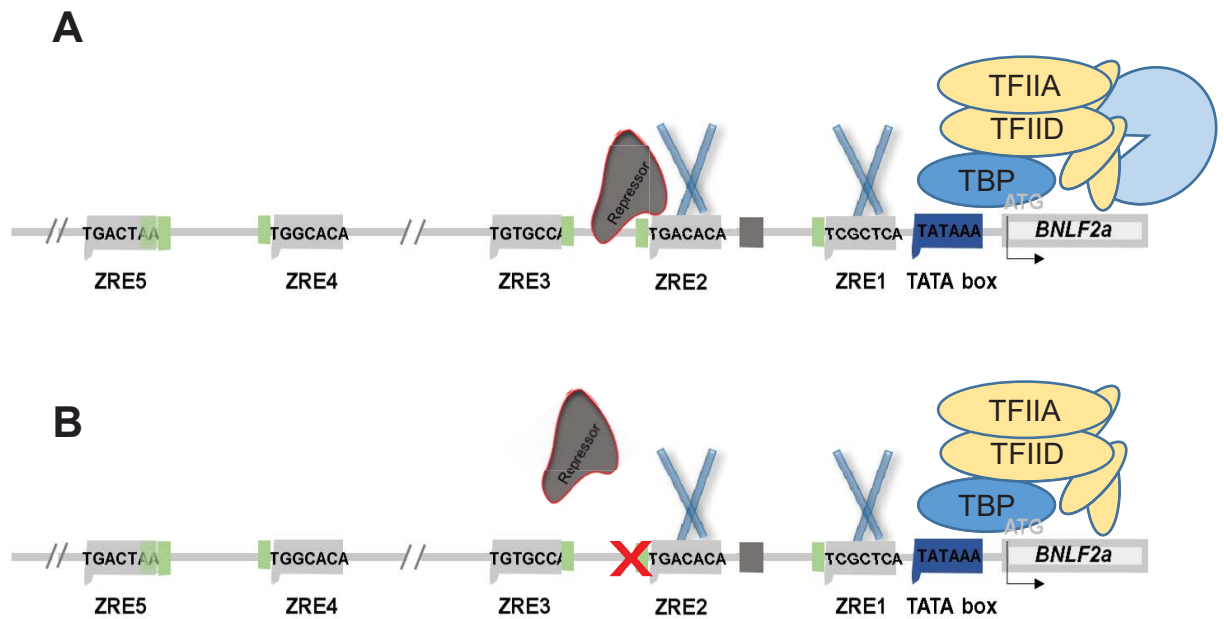
Interestingly, our mutation analysis of this site in HeLa cells has resulted in a sharp increase in the basal activity level in the absence of Zta suggesting a strong role for a repressor protein. The construct BNLF2a 1-5FL M, which contains mutant sequences for all the flanking motifs, has also showed high basal level activity but lower than that seen in ZRE2 flanking mutant. This suggests a further role for other flanking sequences or ZREs in the promoter basal activity level.

Whether the repressor site flanking ZRE2 has a role in Zta activation or not is not clear. We hypothesised that Zta activation is not dependent on this repressor as similar levels of activation were never observed between BNLF2a 1-5 and BNLF2a 2FL M. Also, the relative promoter activity in BNLF2a 2FL M and BNLF2a 1-5FL M was greatly reduced compared to the wildtype BNLF2a 1-5 construct. Thus, it is very likely that the repressor bound at this site is regulated by a different mechanism (**Figure 6-13**).

Based on TOMTOM motif analysis, which allows the comparison of this motif with other position weight matrices of known TFs (Bailey et al., 2009b), two protein TBX and ZEB1 were identified as possible proteins binding to the identified flanking motif (p-value= 1e-4 and p-value= 4.7e-4, respectively).

ZEB1 and ZEB2 are known to be involved in the repression of Zp through E-box sites within the promoter (Ellis et al., 2010) while TBX represents a large family of proteins that share a conserved DNA binding domain and bind to either full or half the T site palindromic motif, which consists of 24 bp (AATTCACACCT-AGGTGTGAAATT) (Najand et al., 2012). Thus, this protein family is more likely to be associated with the longer flanking motif.

Although we showed a specific complex binding to ZRE2 flanking motif in HeLa nuclear extract, this complex is not likely to be ZEB1/2. Thus, further investigation will be required. Another potential target is the E-box binding protein E2-2 that was shown to be involved in Zp silencing (Thomas et al., 2003).



**Figure 6-13 A model for the possible role of ZRE2 flanking motif (E-box).** (A) The ZRE2 flanking motif bound by a repressor does not prevent or compete with Zta binding to ZRE2. Zta is able to overcome the repressing of the promoter. (B) When this site is mutated the repressor is no longer able to bind which results in an increased basal level but this repressor does not stop Zta activation.

## Chapter 7 General discussion

Gene expression regulation takes place at three different levels in higher eukaryotes: the transcriptional, post-transcriptional and post-translational level. To make sure only the set of required genes are expressed at a certain time and in certain cells, this complex process of transcription is tightly regulated temporally and spatially. The synchronous control over gene expression is important to determine cell fate, response to external stimuli (signals) and function. Moreover, alterations in this regulation lead to diseases (Maston et al., 2006).

The initiation of mRNA transcription, which requires the recruitment of the preinitiation complex (PIC), is mostly dependent on transcription factors, regulatory elements, and chromatin accessibility. Transcription factors (TFs) are site-specific DNA molecules that bind to various regulatory elements. These elements include the core and upstream promoters as well as distal enhancers. The core promoter is the site at which the PIC is assembled to execute the initiation of mRNA (Smale and Kadonaga, 2003). The upstream promoter contains multiple cis-elements for the binding of various TFs, which interact with the PIC at the core promoter (Lee and Young, 2000; Maston et al., 2006). Enhancers can be located at 5' or 3' sites and in distal or proximal distances from a given gene. Enhancers play an important role not only to enhance the expression but in many cases to determine tissue specificity of gene expression (Maston et al., 2006).

Given that the genome is highly packed around histones molecules, accessibility to the various elements is not plausible unless these transcription factors work cooperatively to facilitate a more open chromatin state that allows the recruitment of other TFs. A special class of TFs known as the pioneer TFs, have the ability to engage with a nucleosome dense chromatin and recruit histones modifiers (Soufi et al., 2015; Zaret and Carroll, 2011). An example for such pioneer factors is the PU.1 protein, a key regulator in B cell development. This protein is capable of mediating local histone modifications and expanding the linker region between nucleosomes which make it more accessible for other

TFs to bind DNA (Ghisletti et al., 2010). In addition, the key proteins in embryonic stem cell reprogramming Oct4, Sox2 and KLF4 are also classified as pioneer TFs (Ghisletti et al., 2010). Interestingly, the ability of a TF to bind methylated CpG sites has been suggested to be a key feature in characterising pioneer TFs (Zhu et al., 2016). A pioneering TF role was also suggested for the AP-1 protein as this protein was reported to assist in glucocorticoid receptor (GR) binding, by keeping chromatin in an open state (Biddie et al., 2011; Spitz and Furlong, 2012).

Viruses have evolved to hijack and mimic host cell mechanisms to be able to infect and survive within these cells. Upon EBV entry into the cell, the viral DNA becomes circular, and it is packed in higher order repressive chromatin structure, in a very similar way to the host genome. This chromatin state is essential to maintain latency by silencing most of the viral genes (Lieberman, 2013).

Zta which lacks a methyl-CpG binding-domain, has the ability to bind methylated CpG sites. Only a few mammalian proteins are known to bind methylated DNA without methyl-CpG binding-domain, such as the basic leucine zipper (CEBP $\alpha$ ) and the zinc-finger protein ZFP57 (Zhu et al., 2016). Additionally, the AP-1 heterodimer c-Jun/c-Fos was also found to bind a new class of methylated sites (called meAP-1 sites) (Gustems et al., 2014).

The ability of Zta to induce EBV reactivation and reverse the epigenetic repression of the viral genome during latency, through direct binding to methylated ZREs at key viral promoters suggested that Zta may act as a pioneer transcription factor (Woellmer et al., 2012). However, it is not known how Zta is recruited to the densely chromatinised EBV genome and whether Zta has a direct role in chromatin remodeling during lytic cycle (Hammerschmidt, 2015).

As previously discussed in **Chapter 1**, Zta is also essential to recruit the viral DNA replication complex at OriLyt and play a vital role in the regulation of many viral and cellular genes. Zta also interferes with the function of many cellular

pathways that are important to induce apoptosis or host immune response. Zta diverse function strongly suggests a vital pioneering role for this viral transactivator.

To obtain a complete overview of Zta role during EBV lytic cycle, genome-wide studies were carried out previously in our lab and by other groups to detect Zta binding across the viral and cellular genome (Bergbauer et al., 2010; Ramasubramanyan et al., 2012a; Ramasubramanyan et al., 2015).

As previously discussed in **Chapter 3**, I carried out ChIP-qPCR to detect Zta binding across various regions in the viral genome in two different cell lines that are commonly used as a model to study cell transformation and EBV biology. In our experiments, we aimed to address differences observed in Zta binding at different EBV cell lines which will broaden our understanding of the biology and features of each EBV system. For example, we hypothesise that the differences in Zta binding at ~82 kbp region are likely to be caused by EBV strain variations between Raji and B95-8 strains. We also showed Zta binding at several promoter regions in the spontaneously lytic cells (LCL#3). EBV LCLs vary in terms of their surface markers and EBV copy number, which could influence the regulation of the EBV latency-lytic switch (Davies et al., 2010). Similarly, we set out ChIP-qPCR in 293-BZLF1-KO cells (Feederle et al., 2000) in an attempt to understand how Zta binding and the lytic cycle is regulated in a non-B cell system.

To understand the role of Zta in EBV lytic cycle regulation in each EBV cell line, a global approach will be required to investigate Zta binding at the EBV and cellular genome. A thorough analysis of EBV strain variations, EBV copy number, EBV genome chromatin structure, viral and cellular expression profiles will also be crucial. Zta binding data can be integrated with publicly available data for other transcription factors and histone modifications. A similar approach was used to map the binding of various cellular TFs and the chromatin structure organisation in the latent EBV genome (Arvey et al., 2013; Arvey et al., 2012).

The genome-wide analysis for Zta binding across the cellular genome (Ramasubramanian et al., 2015) has also provided a new insight of Zta transcription regulation at enhancer elements. As previously discussed in **Chapter 4**, we showed the first evidence of the ability of Zta to activate a minimal promoter and drive the expression of the luciferase gene in pGL3-Control construct through a distal ZRE element. Zta transcriptional activation through enhancer elements may suggest different regulation mechanisms for Zta that remain to be elucidated.

In this project, we also used the viral promoter *BNLF2a* as a model to broaden our knowledge of Zta transcriptional regulation of viral early lytic genes. As shown in **Chapter 5**, the activation of *BNLF2a* promoter was dependent on three proximal ZREs. Further analysis of these proximal ZREs showed that any combination of two ZREs of the proximal three could drive expression from the promoter. This supports a role for Zta synergy to achieve promoter activation. However, whether the distance between any two ZREs affects this Zta synergy is not clear. Although a previous report showed a general decrease in a luciferase activity when the distance between two ZREs was increased (Huang et al., 2012), the distance between each ZRE and the promoter TSS was not addressed. We hypothesise that if Zta synergy activation is mediated by the simultaneous interaction between the PIC and multiple Zta molecules, the active ZREs should be in close proximity of the core promoter. This also could explain why the distal ZREs in *BNLF2a* promoter did not activate the promoter.

We also addressed the effect of a non-canonical TATA motif (GATA) on Zta activation using *BNLF2a* promoter. Interestingly, the construct containing the GATA motif was around 2-fold higher in relative promoter activity than the wildtype construct. It is not clear whether the presence of such a TATA box variant in the viral gene *BHLF1* would provide any regulation advantages over host or other viral genes. To what extent this TATA variant is present and conserved in other viral or cellular promoters is worth investigating. Furthermore, with the current advances in genome-editing methods, it is possible to generate stable cell lines to investigate the role of GATA motif on the expression of various cellular promoters.

We further attempted to identify and decipher other regulatory elements within *BNLF2a* promoter, as described in **Chapter 6**. *In silico* sequence analysis predicted a 15 bp conserved motif flanking all *BNLF2a* ZREs. Interestingly, when the motif flanking ZRE2 was mutated in our promoter mutational analysis, there was a sharp increase in the basal promoter activity in the absence of Zta which suggests a role for a repressor protein. This flanking motif overlaps a canonical E-box site that was shown before to be important for TPA induction (Jenkins et al., 1997) . A specific DNA-protein complex to this E-box motif was observed in EMSA analysis using HeLa cell nuclear extract. Although we attempted to identify any potential target in this complex, this protein remains to be identified.

Our *BNLF2a* promoter analysis findings open up many interesting questions to address. For example, whether the ZRE2 flanking motif is conserved in other regulatory elements containing ZREs will require genome-wide DNA sequence analysis. This motif may also represent a certain class of promoters or is regulated differently in other cells. Nevertheless, *BNLF2a* regulation may only represent one mechanism of Zta regulation.

Serial deletion analysis of *BNLF2a* promoter will help in identifying the minimal region required for Zta activation. This will also help in identifying other TF binding sites that are required for this activation. On the contrary, Zta is involved in the down regulation of other genes as described earlier. How Zta facilitates the down regulation of such promoters remains to be largely unknown.

Zta activation is also achieved through synergy with other viral proteins such as Rta (BRLF1) and BMRF1. Previous data carried out in our lab by Dr Kay Osborn comparing *BNLF2a* promoter response to Zta in EBV-negative and positive Burkitt's lymphoma cell lines showed similar levels of promoter activation. Although, this may indicate that there is no significant role for other EBV proteins, a synergistic activation between Zta and Rta cannot be ruled out in other cell lines. In fact, Zta and Rta synergy was reported to be more prominent in certain cell lines such as HEK 293 cells (Liu and Speck, 2003). Zta and Rta synergy is widely reported for various lytic cycle promoters such as



*BHLF1*, *BHRF1*, and *BRLF1* (Chang et al., 2010; Liu and Speck, 2003; Ragoczy and Miller, 1999; Yang et al., 2015). Also, it was shown that this synergy is dependent on the direct binding of Zta to the promoter (Chang et al., 2010).

Moreover, it is of great interest to investigate if *BNLF2a* promoter is also responsive to Rta. The immediate early protein Rta is known to transactivate various promoters either directly through its responsive element or indirectly through SP1 sites (Chang et al., 2005). As described in **Chapter 6**, the KLF4 motif within *BNLF2a* promoter is a CC (GG) rich motif which could be bound by Sp1 protein.

The variant mutant constructs generated in this project will be a very valuable tool to look into Zta, Rta, or Zta/Rta transactivation mechanisms. If Zta and Rta synergy is observed in *BNLF2a* promoter, *BNLF2a* various ZRE mutant constructs could be used to identify the sites contributing to this synergy. Also, our purified His-GFP tagged bZIP Zta protein will allow addressing Zta binding to DNA/protein *in vitro* using simple approaches such as EMSA and pull down assays.

Understanding Zta regulation mechanisms is crucial to understand EBV lytic cycle regulation. This will have a wider impact on developing various agents that either inhibit or induce EBV lytic cycle.

## References

- Adams, J. M., A. W. Harris, C. A. Pinkert, L. M. Corcoran, W. S. Alexander, S. Cory, R. D. Palmiter, and R. L. Brinster, 1985, The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice: *Nature*, v. 318, p. 533-8.
- Adamson, A. L., D. Darr, E. Holley-Guthrie, R. A. Johnson, A. Mauser, J. Swenson, and S. Kenney, 2000, Epstein-Barr virus immediate-early proteins BZLF1 and BRLF1 activate the ATF2 transcription factor by increasing the levels of phosphorylated p38 and c-Jun N-terminal kinases: *Journal of Virology*, v. 74, p. 1224-1233.
- Adamson, A. L., and S. Kenney, 1999, The Epstein-Barr virus BZLF1 protein interacts physically and functionally with the histone acetylase CREB-binding protein: *J Virol*, v. 73, p. 6551-8.
- Adamson, A. L., and S. Kenney, 2001, Epstein-barr virus immediate-early protein BZLF1 is SUMO-1 modified and disrupts promyelocytic leukemia bodies: *J Virol*, v. 75, p. 2388-99.
- Adamson, A. L., and S. C. Kenney, 1998, Rescue of the Epstein-Barr virus BZLF1 mutant, Z(S186A), early gene activation defect by the BRLF1 gene product: *Virology*, v. 251, p. 187-197.
- Allan, G. J., and D. T. Rowe, 1989, Size and stability of the Epstein-Barr virus major internal repeat (IR-1) in Burkitt's lymphoma and lymphoblastoid cell lines: *Virology*, v. 173, p. 489-98.
- Allday, M. J., Q. Bazot, and R. E. White, 2015, The EBNA3 Family: Two Oncoproteins and a Tumour Suppressor that Are Central to the Biology of EBV in B Cells: *Curr Top Microbiol Immunol*, v. 391, p. 61-117.
- Allday, M. J., and D. H. Crawford, 1988, Role of epithelium in EBV persistence and pathogenesis of B-cell tumours: *Lancet (London, England)*, v. 1, p. 855-857.
- Allen, M. D., L. S. Young, and C. W. Dawson, 2005, The Epstein-Barr virus-encoded LMP2A and LMP2B proteins promote epithelial cell spreading and motility: *J Virol*, v. 79, p. 1789-1802.
- Alspaugh, M. A., F. C. Jensen, H. Rabin, and E. M. Tan, 1978, Lymphocytes transformed by Epstein-Barr virus. Induction of nuclear antigen reactive with antibody in rheumatoid arthritis: *J Exp Med*, v. 147, p. 1018-27.
- Amon, W., U. K. Binné, H. Bryant, P. J. Jenkins, C. E. Karstegl, and P. J. Farrell, 2004, Lytic cycle gene regulation of Epstein-Barr virus: *Journal of Virology*, v. 78, p. 13460-13469.
- Andtbacka, R. H., H. L. Kaufman, F. Collichio, T. Amatruda, N. Senzer, J. Chesney, K. A. Delman, L. E. Spitler, I. Puzanov, S. S. Agarwala, M. Milhem, L. Cranmer, B. Curti, K. Lewis, M. Ross, T. Guthrie, G. P. Linette, G. A. Daniels, K. Harrington, M. R. Middleton, W. H. Miller, Jr., J. S. Zager, Y. Ye, B. Yao, A. Li, S. Doleman, A. VanderWalde, J. Gansert, and R. S. Coffin, 2015, Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma: *J Clin Oncol*, v. 33, p. 2780-8.
- Arnold, C. D., D. Gerlach, C. Stelzer, L. M. Boryn, M. Rath, and A. Stark, 2013, Genome-wide quantitative enhancer activity maps identified by STARR-seq: *Science*, v. 339, p. 1074-7.
- Arvey, A., I. Tempera, and P. M. Lieberman, 2013, Interpreting the Epstein-Barr Virus (EBV) Epigenome Using High-Throughput Data: *Viruses*, v. 5, p. 1042-1054.
- Arvey, A., I. Tempera, K. Tsai, H.-S. Chen, N. Tikhmyanova, M. Klichinsky, C. Leslie, and P. M. Lieberman, 2012, An Atlas of the Epstein-Barr Virus Transcriptome and Epigenome Reveals Host-Virus Regulatory Interactions: *Cell host & microbe*, v. 12, p. 233-245.

- Arvin, A., G. Campadelli-Fiume, E. Mocarski, P. S. Moore, B. Roizman, R. Whitley, and K. Yamanishi, 2007, Human herpesviruses : biology, therapy, and immunoprophylaxis / edited by Ann Arvin ... [et al.]. Cambridge ; New York, Cambridge University Press, 1388 p p.
- Asai, R., A. Kato, K. Kato, M. Kanamori-Koyama, K. Sugimoto, T. Sairenji, Y. Nishiyama, and Y. Kawaguchi, 2006, Epstein-Barr virus protein kinase BGLF4 is a virion tegument protein that dissociates from virions in a phosphorylation-dependent process and phosphorylates the viral immediate-early protein BZLF1: *Journal of Virology*, v. 80, p. 5125-5134.
- Asai, R., A. Kato, and Y. Kawaguchi, 2009, Epstein-Barr virus protein kinase BGLF4 interacts with viral transactivator BZLF1 and regulates its transactivation activity: *J Gen Virol*, v. 90, p. 1575-81.
- Ascherio, A., and K. L. Munger, 2015, EBV and Autoimmunity: *Curr Top Microbiol Immunol*, v. 390, p. 365-85.
- Asito, A. S., E. Piriou, P. S. Odada, N. Fiore, J. M. Middeldorp, C. Long, S. Dutta, D. E. Lanar, W. G. Jura, C. Ouma, J. A. Otieno, A. M. Moormann, and R. Rochford, 2010, Elevated anti-Zta IgG levels and EBV viral load are associated with site of tumor presentation in endemic Burkitt's lymphoma patients: a case control study: *Infectious Agents and Cancer*, v. 5, p. 13-13.
- Askovic, S., and R. Baumann, 1997, Activation domain requirements for disruption of Epstein-Barr virus latency by ZEBRA: *J Virol*, v. 71, p. 6547-54.
- Atchley, W. R., and W. M. Fitch, 1997, A natural classification of the basic helix-loop-helix class of transcription factors: *Proc Natl Acad Sci U S A*, v. 94, p. 5172-6.
- Aubry, V., F. Mure, B. Mariamé, T. Deschamps, L. S. Wyrwicz, E. Manet, and H. Gruffat, 2014, Epstein-Barr Virus Late Gene Transcription Depends on the Assembly of a Virus-Specific Preinitiation Complex: *Journal of Virology*, v. 88, p. 12825-12838.
- Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Séguin, P. S. Tuffnell, and B. G. Barrell, 1984, DNA sequence and expression of the B95-8 Epstein - Barr virus genome: *Nature*, v. 310, p. 207-211.
- Bailey, S. G., E. Verrall, C. Schelcher, A. Rhie, A. J. Doherty, and A. J. Sinclair, 2009a, Functional interaction between Epstein-Barr virus replication protein Zta and host DNA damage response protein 53BP1: *J Virol*, v. 83, p. 11116-22.
- Bailey, T. L., M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Ren, W. W. Li, and W. S. Noble, 2009b, MEME SUITE: tools for motif discovery and searching: *Nucleic Acids Res*, v. 37, p. W202-8.
- Balan, N., K. Osborn, and A. J. Sinclair, 2016, Repression of CIITA by the Epstein-Barr virus transcription factor Zta is independent of its dimerization and DNA binding: *J Gen Virol*, v. 97, p. 725-32.
- Balfour, H. H., O. A. Odumade, D. O. Schmeling, B. D. Mullan, J. A. Ed, J. A. Knight, H. E. Vezina, W. Thomas, and K. A. Hogquist, 2013, Behavioral, Virologic, and Immunologic Factors Associated With Acquisition and Severity of Primary Epstein-Barr Virus Infection in University Students: *The Journal of Infectious Diseases*, v. 207, p. 80-88.
- Banerji, J., S. Rusconi, and W. Schaffner, 1981, Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences: *Cell*, v. 27, p. 299-308.
- Barth, S., T. Pfuhl, A. Mamiani, C. Ehse, K. Roemer, and E. Kremmer, 2008, Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5: *Nucleic Acids Res*, v. 36, p. 666-675.
- Becker, J., U. Leser, M. Marschall, A. Langford, W. Jilg, H. Gelderblom, P. Reichart, and H. Wolf, 1991, Expression of proteins encoded by Epstein-Barr virus trans-activator genes depends on the differentiation of epithelial cells in oral hairy leukoplakia: *Proceedings of the National Academy of Sciences of the United States of America*, v. 88, p. 8332-8336.

- Bell, M. J., R. J. Abbott, N. P. Croft, A. D. Hislop, and S. R. Burrows, 2009, An HLA-A2-restricted T-cell epitope mapped to the BNLF2a immune evasion protein of Epstein-Barr virus that inhibits TAP: *J Virol*, v. 83, p. 2783-8.
- Bell, P., P. M. Lieberman, and G. G. Maul, 2000, Lytic but not latent replication of Epstein-Barr virus is associated with PML and induces sequential release of nuclear domain 10 proteins: *Journal of Virology*, v. 74, p. 11800-11810.
- Ben-Bassat, H., N. Goldblum, S. Mitrani, T. Goldblum, J. M. Yoffey, M. M. Cohen, Z. Bentwich, B. Ramot, E. Klein, and G. Klein, 1977, Establishment in continuous culture of a new type of lymphocyte from a "Burkitt like" malignant lymphoma (line D.G.-75): *Int J Cancer*, v. 19, p. 27-33.
- Bergbauer, M., M. Kalla, A. Schmeinck, C. Göbel, U. Rothbauer, S. Eck, A. Benet-Pagés, T. M. Strom, and W. Hammerschmidt, 2010, CpG-methylation regulates a class of Epstein-Barr virus promoters: *PLoS Pathogens*, v. 6.
- Berman, B. P., Y. Nibu, B. D. Pfeiffer, P. Tomancak, S. E. Celniker, M. Levine, G. M. Rubin, and M. B. Eisen, 2002, Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the *Drosophila* genome: *Proc Natl Acad Sci U S A*, v. 99, p. 757-62.
- Bernig, T., N. Richter, I. Volkmer, and M. S. Staeger, 2014, Functional analysis and molecular characterization of spontaneously outgrown human lymphoblastoid cell lines: *Molecular Biology Reports*, v. 41, p. 6995-7007.
- Beswick, T. S. L., 1962, The origin and the use of the word herpes: *Medical History*, v. 6, p. 214-232.
- Bhende, P. M., S. J. Dickerson, X. Sun, W. H. Feng, and S. C. Kenney, 2007, X-box-binding protein 1 activates lytic Epstein-Barr virus gene expression in combination with protein kinase D: *Journal of Virology*, v. 81, p. 7363-7370.
- Bhende, P. M., W. T. Seaman, H. J. Delecluse, and S. C. Kenney, 2004, The EBV lytic switch protein, Z, preferentially binds to and activates the methylated viral genome: *Nat Genet*, v. 36, p. 1099-104.
- Bhende, P. M., W. T. Seaman, H. J. Delecluse, and S. C. Kenney, 2005, BZLF1 activation of the methylated form of the BRLF1 immediate-early promoter is regulated by BZLF1 residue 186: *J Virol*, v. 79, p. 7338-48.
- Biddie, Simon C., S. John, Pete J. Sabo, Robert E. Thurman, Thomas A. Johnson, R. L. Schiltz, Tina B. Miranda, M.-H. Sung, S. Trump, Stafford L. Lightman, C. Vinson, John A. Stamatoyannopoulos, and Gordon L. Hager, 2011, Transcription Factor AP1 Potentiates Chromatin Accessibility and Glucocorticoid Receptor Binding: *Molecular Cell*, v. 43, p. 145-155.
- Blacklow, N. R., B. K. Watson, G. Miller, and B. M. Jacobson, 1971, Mononucleosis with heterophil antibodies and EB virus infection: *The American Journal of Medicine*, v. 51, p. 549-552.
- Bogedain, C., P. Alliger, F. Schwarzmann, M. Marschall, H. Wolf, and W. Jilg, 1994, Different activation of Epstein-Barr virus immediate-early and early genes in Burkitt lymphoma cells and lymphoblastoid cell lines: *Journal of Virology*, v. 68, p. 1200-1203.
- Borer, A., J. Gilad, H. Haikin, K. Riesenber, A. Porath, and F. Schlaeffer, 1999, Clinical features and costs of care for hospitalized adults with primary Epstein-Barr virus infection: *Am J Med*, v. 107, p. 144-8.
- Bornkamm, G. W., and W. Hammerschmidt, 2001, Molecular virology of Epstein-Barr virus: *Philos Trans R Soc Lond B Biol Sci*, v. 356, p. 437-59.
- Brady, G., G. J. MacArthur, and P. J. Farrell, 2007, Epstein-Barr virus and Burkitt lymphoma: *Journal of Clinical Pathology*, v. 60, p. 1397-1402.
- Brubaker, S. W., K. S. Bonham, I. Zanoni, and J. C. Kagan, 2015, Innate immune pattern recognition: a cell biological perspective: *Annu Rev Immunol*, v. 33, p. 257-90.

- Bryant, H., and P. J. Farrell, 2002, Signal Transduction and Transcription Factor Modification during Reactivation of Epstein-Barr Virus from Latency: *J Virol*, v. 76, p. 10290-8.
- Buell, P., 1974, The effect of migration on the risk of nasopharyngeal cancer among Chinese: *Cancer Res*, v. 34, p. 1189-91.
- Buisson, M., E. Manet, M. C. Trescol-Biemont, H. Gruffat, B. Durand, and A. Sergeant, 1989, The Epstein-Barr virus (EBV) early protein EB2 is a posttranscriptional activator expressed under the control of EBV transcription factors EB1 and R: *Journal of Virology*, v. 63, p. 5276-5284.
- Burkitt, D., 1958, A sarcoma involving the jaws in African children: *Br J Surg*, v. 46, p. 218-23.
- Burkitt, D., 1962a, A "tumour safari" in East and Central Africa: *Br J Cancer*, v. 16, p. 379-86.
- Burkitt, D., 1962b, A children's cancer dependent on climatic factors: *Nature*, v. 194, p. 232-4.
- Burkitt, D., 1965, Chemotherapy of jaw lymphomata: *East Afr Med J*, v. 42, p. 244-8.
- Burkitt, D., M. S. Hutt, and D. H. Wright, 1965, The African lymphoma: preliminary observations on response to therapy: *Cancer*, v. 18, p. 399-410.
- Burkitt, D., and G. T. O'Connor, 1961, Malignant lymphoma in African children. I. A clinical syndrome: *Cancer*, v. 14, p. 258-69.
- Burkitt, D., and D. H. Wright, 1963, A lymphoma syndrome in tropical Africa with anote on histology, cytology, and histochemistry: *Int Rev Exp Pathol*, v. 2, p. 67-138.
- Burkitt, D. P., 1961, Observations on the geography of malignant lymphoma: *East Afr Med J*, v. 38, p. 511-4.
- Butka, H. E., 1926, Infectious mononucleosis: With report of five cases: *California and Western Medicine*, v. 25, p. 353-355.
- Cai, Q., K. Chen, and K. H. Young, 2015, Epstein-Barr virus-positive T/NK-cell lymphoproliferative disorders: *Exp Mol Med*, v. 47, p. e133.
- Cai, X., A. Schafer, S. Lu, J. P. Bilello, R. C. Desrosiers, and R. Edwards, 2006, Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed: *PLoS Pathogen*, v. 2, p. e23.
- Cameron, K. R., T. Stamminger, M. Craxton, W. Bodemer, R. W. Honess, and B. Fleckenstein, 1987, The 160,000-Mr virion protein encoded at the right end of the herpesvirus saimiri genome is homologous to the 140,000-Mr membrane antigen encoded at the left end of the Epstein-Barr virus genome: *J Virol*, v. 61, p. 2063-70.
- Carbone, A., A. Gloghini, and G. Dotti, 2008, EBV-associated lymphoproliferative disorders: classification and treatment: *Oncologist*, v. 13, p. 577-85.
- Cardiff, R. D., and N. Kenney, 2007, Mouse mammary tumor biology: A short history, *Advances in Cancer Research*, v. Volume 98, Academic Press, p. 53-116.
- Carey, M., J. Kolman, D. A. Katz, L. Gradoville, L. Barberis, and G. Miller, 1992, Transcriptional synergy by the Epstein-Barr virus transactivator ZEBRA: *J Virol*, v. 66, p. 4803-13.
- Cayrol, C., and E. Flemington, 1996a, G0/G1 growth arrest mediated by a region encompassing the basic leucine zipper (bZIP) domain of the Epstein-Barr virus transactivator Zta: *Journal of Biological Chemistry*, v. 271, p. 31799-31802.
- Cayrol, C., and E. K. Flemington, 1996b, The Epstein-Barr virus bZIP transcription factor Zta causes G0/G1 cell cycle arrest through induction of cyclin-dependent kinase inhibitors: *EMBO J*, v. 15, p. 2748-59.
- Cen, O., and R. Longnecker, 2015, Latent Membrane Protein 2 (LMP2): *Curr Top Microbiol Immunol*, v. 391, p. 151-80.
- Chang, E. T., and H. O. Adami, 2006, The enigmatic epidemiology of nasopharyngeal carcinoma: *Cancer Epidemiol Biomarkers Prev*, v. 15, p. 1765-77.



- Chang, L. K., J. Y. Chuang, M. Nakao, and S. T. Liu, 2010, MCAF1 and synergistic activation of the transcription of Epstein-Barr virus lytic genes by Rta and Zta: *Nucleic Acids Res*, v. 38, p. 4687-700.
- Chang, L. K., J. Y. Chung, Y. R. Hong, T. Ichimura, M. Nakao, and S. T. Liu, 2005, Activation of Sp1-mediated transcription by Rta of Epstein-Barr virus via an interaction with MCAF1: *Nucleic Acids Res*, v. 33, p. 6528-39.
- Chang, Y., D. Dong, G. Hayward, and S. D. Hayward, 1990, The Epstein-Barr virus Zta transactivator: a member of bZIP family with unique DNA-binding specificity and a dimerization domain that lacks the characteristic heptad leucine zipper motif: *J. Virol.*, v. 64, p. 3358-3369.
- Chang, Y., H. H. Lee, Y. T. Chen, J. Lu, S. Y. Wu, C. W. Chen, K. Takada, and C. H. Tsai, 2006, Induction of the early growth response 1 gene by Epstein-Barr virus lytic transactivator Zta: *J Virol*, v. 80, p. 7748-55.
- Chen, C., D. Li, and N. Guo, 2009, Regulation of cellular and viral protein expression by the Epstein-Barr virus transcriptional regulator Zta: Implications for therapy of EBV associated tumors: *Cancer Biology & Therapy*, v. 8, p. 987-995.
- Chen, C. C., Y. C. Yang, W. H. Wang, C. S. Chen, and L. K. Chang, 2011, Enhancement of Zta-activated lytic transcription of Epstein-Barr virus by Ku80: *J Gen Virol*, v. 92, p. 661-8.
- Chen, C. J., Z. Deng, A. Y. Kim, G. A. Blobel, and P. M. Lieberman, 2001, Stimulation of CREB binding protein nucleosomal histone acetyltransferase activity by a class of transcriptional activators: *Mol Cell Biol*, v. 21, p. 476-87.
- Chen, H., J. Huang, F. Y. Wu, G. Liao, L. Hutt-Fletcher, and S. D. Hayward, 2005, Regulation of expression of the Epstein-Barr virus BamHI-a rightward transcripts: *J Virol*, v. 79.
- Chen, H., P. Smith, R. F. Ambinder, and S. D. Hayward, 1999, Expression of Epstein-Barr virus BamHI-A rightward transcripts in latently infected B cells from peripheral blood: *Blood*, v. 93, p. 3026-3032.
- Chen, H. S., F. Lu, and P. M. Lieberman, 2013, Epigenetic regulation of EBV and KSHV latency: *Curr Opin Virol*, v. 3, p. 251-9.
- Chen, M. R., 2011, Epstein-barr virus, the immune system, and associated diseases: *Front Microbiol*, v. 2, p. 5.
- Chen, M. R., S. J. Chang, H. Huang, and J. Y. Chen, 2000, A protein kinase activity associated with Epstein-Barr virus BGLF4 phosphorylates the viral early antigen EA-D in vitro: *Journal of Virology*, v. 74, p. 3093-3104.
- Chen, Y., W. Zhao, L. Lin, X. Xiao, X. Zhou, H. Ming, T. Huang, J. Liao, Y. Li, X. Zeng, G. Huang, W. Ye, and Z. Zhang, 2015, Nasopharyngeal Epstein-Barr Virus Load: An Efficient Supplementary Method for Population-Based Nasopharyngeal Carcinoma Screening: *PLoS ONE*, v. 10, p. e0132669.
- Chesnokova, L. S., and L. M. Hutt-Fletcher, 2014, Epstein-Barr virus infection mechanisms: *Chinese Journal of Cancer*, v. 33, p. 545-548.
- Chesnokova, L. S., R. Jiang, and L. M. Hutt-Fletcher, 2015, Viral Entry, *in* C. Münz, ed., *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 221-235.
- Cheung, A., and E. Kieff, 1982, Long internal direct repeat in Epstein-Barr virus DNA: *Journal of Virology*, v. 44, p. 286-294.
- Chevallier-Greco, A., E. Manet, P. Chavrier, C. Mosnier, J. Daillie, and A. Sergeant, 1986, Both Epstein-Barr virus (EBV)-encoded trans-acting factors, EB1 and EB2, are required to activate transcription from an EBV early promoter: *Embo j*, v. 5, p. 3243-9.
- Chi, T., and M. Carey, 1993, The ZEBRA activation domain: Modular organization and mechanism of action: *Molecular and Cellular Biology*, v. 13, p. 7045-7055.
- Cho, J., M.-S. Kang, and K.-M. Kim, 2016, Epstein-Barr virus-associated gastric carcinoma and specific features of the accompanying immune response: *Journal of Gastric Cancer*, v. 16, p. 1-7.

- Cho, M. S., G. W. Bornkamm, and H. zur Hausen, 1984, Structure of defective DNA molecules in Epstein-Barr virus preparations from P3HR-1 cells: *J Virol*, v. 51, p. 199-207.
- Cohen, J. I., 2000, Epstein-Barr virus infection: *N Engl J Med*, v. 343, p. 481-92.
- Costa, L. J., A. C. Xavier, A. E. Wahlquist, and E. G. Hill, 2013, Trends in survival of patients with Burkitt lymphoma/leukemia in the USA: an analysis of 3691 cases: *Blood*, v. 121, p. 4861-4866.
- Costenbader, K. H., and E. W. Karlson, 2006, Epstein-Barr virus and rheumatoid arthritis: is there a link?: *Arthritis Research & Therapy*, v. 8, p. 204-204.
- Countryman, J., and G. Miller, 1985, Activation of expression of latent Epstein-Barr herpesvirus after gene transfer with a small cloned subfragment of heterogeneous viral DNA: *Proceedings of the National Academy of Sciences of the United States of America*, v. 82, p. 4085-4089.
- Cox, M. A., J. Leahy, and J. Marie Hardwick, 1990, An enhancer within the divergent promoter of Epstein-Barr virus responds synergistically to the R and Z transactivators: *Journal of Virology*, v. 64, p. 313-321.
- Crawford, D. H., and I. Ando, 1986, EB virus induction is associated with B-cell maturation: *Immunology*, v. 59, p. 405-9.
- Crawford, D. H. R. A., Johannessen Ingólfur., 2014, *Cancer Virus*: Oxford press.
- Croft, N. P., C. Shannon-Lowe, A. I. Bell, D. Horst, E. Kremmer, M. E. Rensing, E. J. Wiertz, J. M. Middeldorp, M. Rowe, A. B. Rickinson, and A. D. Hislop, 2009, Stage-specific inhibition of MHC class I presentation by the Epstein-Barr virus BNLF2a protein during virus lytic cycle: *PLoS Pathog*, v. 5, p. e1000490.
- Crosignani, P., A. De Stefani, G. M. Fara, A. M. Isidori, A. Lenzi, C. A. Liverani, A. Lombardi, F. S. Mennini, G. Palu, S. Pecorelli, A. P. Peracino, C. Signorelli, and G. V. Zuccotti, 2013, Towards the eradication of HPV infection through universal specific vaccination: *BMC Public Health*, v. 13, p. 642.
- Dalla-Favera, R., M. Bregni, J. Erikson, D. Patterson, R. C. Gallo, and C. M. Croce, 1982, Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells: *Proc Natl Acad Sci U S A*, v. 79, p. 7824-7.
- Dalldorf, G., C. A. Linsell, F. E. Barnhart, and R. Martyn, 1964, An epidemiologic approach to the lymphomas of African children and Burkitt's sarcoma of the jaws: *Perspect Biol Med*, v. 7, p. 435-49.
- Dambaugh, T. R., and E. Kieff, 1982, Identification and nucleotide sequences of two similar tandem direct repeats in Epstein-Barr virus DNA: *J Virol*, v. 44, p. 823-33.
- Dantuma, N. P., A. Sharipo, and M. G. Masucci, 2002, Avoiding proteasomal processing: the case of EBNA1: *Curr Top Microbiol Immunol*, v. 269, p. 23-36.
- Darr, C. D., A. Mauser, and S. Kenney, 2001, Epstein-Barr virus immediate-early protein BRLF1 induces the lytic form of viral replication through a mechanism involving phosphatidylinositol-3 kinase activation: *J Virol*, v. 75, p. 6135-42.
- Datta, A. K., B. M. Colby, J. E. Shaw, and J. S. Pagano, 1980, Acyclovir inhibition of Epstein-Barr virus replication: *Proceedings of the National Academy of Sciences of the United States of America*, v. 77, p. 5163-5166.
- Davies, M. L., S. Xu, J. Lyons-Weiler, A. Rosendorff, S. A. Webber, L. R. Wasil, D. Metes, and D. T. Rowe, 2010, Cellular factors associated with latency and spontaneous Epstein-Barr virus reactivation in B-lymphoblastoid cell lines: *Virology*, v. 400, p. 53-67.
- Davison, A. J., 2010, Herpesvirus systematics: *Vet Microbiol*, v. 143, p. 52-69.
- Davison, A. J., R. Eberle, B. Ehlers, G. S. Hayward, D. J. McGeoch, A. C. Minson, P. E. Pellett, B. Roizman, M. J. Studdert, and E. Thiry, 2009, The Order Herpesvirales: *Archives of virology*, v. 154, p. 171-177.
- De Jesus, N. H., 2007, Epidemics to eradication: the modern history of poliomyelitis: *Virology Journal*, v. 4, p. 1-18.

- de Jesus, O., P. R. Smith, L. C. Spender, C. Elgueta Karstegl, H. H. Niller, D. Huang, and P. J. Farrell, 2003, Updated Epstein-Barr virus (EBV) DNA sequence and analysis of a promoter for the BART (CST, BARF0) RNAs of EBV: *J Gen Virol*, v. 84, p. 1443-50.
- de Martel, C., J. Ferlay, S. Franceschi, J. Vignat, F. Bray, D. Forman, and M. Plummer, 2012, Global burden of cancers attributable to infections in 2008: a review and synthetic analysis: *The Lancet Oncology*, v. 13, p. 607-615.
- de Wet, J. R., K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani, 1987, Firefly luciferase gene: structure and expression in mammalian cells: *Molecular and Cellular Biology*, v. 7, p. 725-737.
- de-The, G., N. E. Day, A. Geser, M. F. Lavoue, J. H. Ho, M. J. Simons, R. Sohier, P. Tukei, V. Vonka, and H. Zavadova, 1975, Sero-epidemiology of the Epstein-Barr virus: preliminary analysis of an international study - a review: *IARC Sci Publ*, p. 3-16.
- Decaussin, G., V. Leclerc, and T. Ooka, 1995, The lytic cycle of Epstein-Barr virus in the nonproducer Raji line can be rescued by the expression of a 135-kilodalton protein encoded by the BALF2 open reading frame: *J Virol*, v. 69, p. 7309-14.
- Deng, Z., 2003, The CBP bromodomain and nucleosome targeting are required for Zta-directed nucleosome acetylation and transcription activation: *Mol. Cell. Biol.*, v. 23, p. 2633-2644.
- Deng, Z., C.-J. Chen, D. Zerby, H.-J. Delecluse, and P. M. Lieberman, 2001, Identification of Acidic and Aromatic Residues in the Zta Activation Domain Essential for Epstein-Barr Virus Reactivation: *Journal of Virology*, v. 75, p. 10334-10347.
- Diala, E. S., and R. M. Hoffman, 1983, Epstein-Barr HR-1 Virion DNA Is Very Highly Methylated: *Journal of Virology*, v. 45, p. 482-483.
- Dickerson, S. J., Y. Xing, A. R. Robinson, W. T. Seaman, H. Gruffat, and S. C. Kenney, 2009, Methylation-dependent binding of the Epstein-Barr virus BZLF1 protein to viral promoters: *PLoS Pathog*, v. 5, p. e1000356.
- Dierickx, D., T. Tousseyn, and O. Gheysens, 2015, How I treat posttransplant lymphoproliferative disorders: *Blood*, v. 126, p. 2274.
- Dolan, A., C. Addison, D. Gatherer, A. J. Davison, and D. J. McGeoch, 2006, The genome of Epstein-Barr virus type 2 strain AG876: *Virology*, v. 350, p. 164-70.
- Dreyfus, D. H., Y. Liu, L. Y. Ghoda, and J. T. Chang, 2011, Analysis of an ankyrin-like region in Epstein Barr Virus encoded (EBV) BZLF-1 (ZEBRA) protein: implications for interactions with NF- $\kappa$ B and p53: *Virology Journal*, v. 8, p. 1-17.
- Dunmire, S. K., K. A. Hogquist, and H. H. Balfour, 2015, Infectious Mononucleosis: Current topics in microbiology and immunology, v. 390, p. 211-240.
- Dyson, P. J., and P. J. Farrell, 1985, Chromatin structure of Epstein-Barr virus: *J Gen Virol*, v. 66 ( Pt 9), p. 1931-40.
- El-Guindy, A., M. Ghiassi-Nejad, S. Golden, H. J. Delecluse, and G. Miller, 2013, Essential role of Rta in lytic DNA replication of Epstein-Barr virus: *J Virol*, v. 87, p. 208-23.
- El-Guindy, A., L. Heston, H. J. Delecluse, and G. Miller, 2007, Phosphoacceptor site S173 in the regulatory domain of Epstein-Barr virus ZEBRA protein is required for lytic DNA replication but not for activation of viral early genes: *Journal of Virology*, v. 81, p. 3303-3316.
- El-Guindy, A. S., and G. Miller, 2004, Phosphorylation of Epstein-Barr virus ZEBRA protein at its casein kinase 2 sites mediates its ability to repress activation of a viral lytic cycle late gene by Rta: *Journal of Virology*, v. 78, p. 7634-7644.
- Ellis, A. L., Z. Wang, X. Yu, and J. E. Mertz, 2010, Either ZEB1 or ZEB2/SIP1 can play a central role in regulating the Epstein-Barr virus latent-lytic switch in a cell-type-specific manner: *J Virol*, v. 84, p. 6139-52.



- Emens, L. A., L. H. Butterfield, F. S. Hodi, F. M. Marincola, and H. L. Kaufman, 2016, Cancer immunotherapy trials: leading a paradigm shift in drug development: *Journal for Immunotherapy of Cancer*, v. 4, p. 42.
- Epstein, A., 2015, Why and How Epstein-Barr Virus Was Discovered 50 Years Ago, *in* C. Münz, ed., *Epstein Barr Virus Volume 1: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 3-15.
- Epstein, M. A., 2001, Historical background: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, v. 356, p. 413-420.
- Epstein, M. A., G. Henle, B. G. Achong, and Y. M. Barr, 1965, Morphological and biological studies on a virus in cultured lymphoblasts from Burkitt's lymphoma: *The Journal of Experimental Medicine*, v. 121, p. 761-770.
- Epstein, M. A., and S. J. Holt, 1958, Observations on the Rous Virus; Integrated Electron Microscopical and Cytochemical Studies of Fluorocarbon Purified Preparations: *British Journal of Cancer*, v. 12, p. 363-369.
- Evans, A. S., N. F. Rothfield, and J. C. Niederman, 1971, Raised antibody titres to E.B. virus in systemic lupus erythematosus: *Lancet*, v. 1, p. 167-8.
- Farre, D., R. Roset, M. Huerta, J. E. Adsua, L. Rosello, M. M. Alba, and X. Messegue, 2003, Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN: *Nucleic Acids Res*, v. 31, p. 3651-3.
- Farrell, P. J., 2001, Epstein-Barr virus. The B95-8 strain map: *Methods Mol Biol*, v. 174, p. 3-12.
- Farrell, P. J., 2015, Epstein-Barr Virus Strain Variation, *in* C. Münz, ed., *Epstein Barr Virus Volume 1: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 45-69.
- Farrell, P. J., A. Bankier, C. Séguin, P. Deininger, and B. G. Barrell, 1983, Latent and lytic cycle promoters of Epstein-Barr virus: *The EMBO Journal*, v. 2, p. 1331-1338.
- Farrell, P. J., D. T. Rowe, C. M. Rooney, and T. Kouzarides, 1989, Epstein-Barr virus BZLF1 trans-activator specifically binds to a consensus AP-1 site and is related to c-fos: *Embo j*, v. 8, p. 127-32.
- Feederle, R., E. J. Bartlett, and H. J. Delecluse, 2010, Epstein-Barr virus genetics: talking about the BAC generation: *Herpesviridae*, v. 1, p. 6.
- Feederle, R., O. Klinke, A. Kutikhin, R. Poirey, M. H. Tsai, and H. J. Delecluse, 2015, Epstein-Barr Virus: From the Detection of Sequence Polymorphisms to the Recognition of Viral Types: *Curr Top Microbiol Immunol*, v. 390, p. 119-48.
- Feederle, R., M. Kost, M. Baumann, A. Janz, E. Drouet, W. Hammerschmidt, and H. J. Delecluse, 2000, The Epstein-Barr virus lytic program is controlled by the co-operative functions of two transactivators: *EMBO J*, v. 19, p. 3080-9.
- Feederle, R., B. Neuhierl, H. Bannert, K. Geletneky, C. Shannon-Lowe, and H. J. Delecluse, 2007, Epstein-Barr virus B95.8 produced in 293 cells shows marked tropism for differentiated primary epithelial cells and reveals interindividual variation in susceptibility to viral infection: *Int J Cancer*, v. 121, p. 588-94.
- Feng, P., E. C. Ren, D. Liu, S. H. Chan, and H. Hu, 2000, Expression of Epstein-Barr virus lytic gene BRLF1 in nasopharyngeal carcinoma: potential use in diagnosis: *Journal of General Virology*, v. 81, p. 2417-2423.
- Fields, B. N., D. M. Knipe, and P. M. Howley, 2013, *Fields virology*: Philadelphia, Wolters Kluwer Health/Lippincott Williams & Wilkins.
- Fingerroth, J. D., M. E. Diamond, D. R. Sage, J. Hayman, and J. L. Yates, 1999, CD21-Dependent infection of an epithelial cell line, 293, by Epstein-Barr virus: *J Virol*, v. 73, p. 2115-25.
- Fingerroth, J. D., J. J. Weis, T. F. Tedder, J. L. Strominger, P. A. Biro, and D. T. Fearon, 1984, Epstein-Barr virus receptor of human B lymphocytes is the C3d receptor CR2: *Proc Natl Acad Sci USA*, v. 81, p. 4510-4514.
- Fixman, E. D., G. S. Hayward, and S. D. Hayward, 1992, Trans-acting requirements for replication of Epstein-Barr virus ori-Lyt: *Journal of Virology*, v. 66, p. 5030-5039.

- Fixman, E. D., G. S. Hayward, and S. D. Hayward, 1995, Replication of Epstein-Barr virus oriLyf: lack of a dedicated virally encoded origin-binding protein and dependence on Zta in cotransfection assays: *J Virol*, v. 69, p. 2998-3006.
- Fleisher, G., W. Henle, G. Henle, E. T. Lennette, and R. J. Biggar, 1979, Primary infection with Epstein-Barr virus in infants in the United States: clinical and serologic observations: *J Infect Dis*, v. 139, p. 553-8.
- Flemington, E., and S. H. Speck, 1990a, Autoregulation of Epstein-Barr virus putative lytic switch gene BZLF1: *J Virol*, v. 64, p. 1227-32.
- Flemington, E., and S. H. Speck, 1990b, Epstein-Barr virus BZLF1 trans activator induces the promoter of a cellular cognate gene, c-fos: *Journal of Virology*, v. 64, p. 4549-4552.
- Flemington, E., and S. H. Speck, 1990c, Evidence for coiled-coil dimer formation by an Epstein-Barr virus transactivator that lacks a heptad repeat of leucine residues: *Proceedings of the National Academy of Sciences of the United States of America*, v. 87, p. 9459-9463.
- Flemington, E. K., A. M. Borrás, J. P. Lytle, and S. H. Speck, 1992, Characterization of the Epstein-Barr virus BZLF1 protein transactivation domain: *J Virol*, v. 66, p. 922-9.
- Flower, K., D. Thomas, J. Heather, S. Ramasubramanyan, S. Jones, and A. J. Sinclair, 2011, Epigenetic Control of Viral Life-Cycle by a DNA-Methylation Dependent Transcription Factor: *PLoS One*, v. 6.
- Fong, T. C., and B. M. Emerson, 1992, The erythroid-specific protein cGATA-1 mediates distal enhancer activity through a specialized beta-globin TATA box: *Genes Dev*, v. 6, p. 521-32.
- Francis, A. L., L. Gradoville, and G. Miller, 1997, Alteration of a single serine in the basic domain of the Epstein-Barr virus ZEBRA protein separates its functions of transcriptional activation and disruption of latency: *Journal of Virology*, v. 71, p. 3054-3061.
- Frapplier, L., 2015, EBNA1, in C. Münz, ed., *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 3-34.
- Funnell, A. P., L. J. Norton, K. S. Mak, J. Burdach, C. M. Artuz, N. A. Twine, M. R. Wilkins, C. A. Power, T. T. Hung, J. Perdomo, P. Koh, K. S. Bell-Anderson, S. H. Orkin, S. T. Fraser, A. C. Perkins, R. C. Pearson, and M. Crossley, 2012, The CACCC-binding protein KLF3/BKLF represses a subset of KLF1/EKLF target genes and is required for proper erythroid maturation in vivo: *Mol Cell Biol*, v. 32, p. 3281-92.
- Gade, P., and D. V. Kalvakolanu, 2012, Chromatin Immunoprecipitation Assay as a Tool for Analyzing Transcription Factor Activity: *Methods in molecular biology* (Clifton, N.J.), v. 809, p. 85-104.
- Garrett-Sinha, L. A., H. Eberspaecher, M. F. Seldin, and B. de Crombrughe, 1996, A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells: *J Biol Chem*, v. 271, p. 31384-90.
- George, M. R., 2014, Hemophagocytic lymphohistiocytosis: review of etiologies and management: *Journal of Blood Medicine*, v. 5, p. 69-86.
- Germain, R. N., 1994, MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation: *Cell*, v. 76, p. 287-99.
- Gerstein, M. B., A. Kundaje, M. Hariharan, S. G. Landt, K. K. Yan, C. Cheng, X. J. Mu, E. Khurana, J. Rozowsky, R. Alexander, R. Min, P. Alves, A. Abyzov, N. Addleman, N. Bhardwaj, A. P. Boyle, P. Cayting, A. Charos, D. Z. Chen, Y. Cheng, D. Clarke, C. Eastman, G. Euskirchen, S. Fietze, Y. Fu, J. Gertz, F. Grubert, A. Harmanci, P. Jain, M. Kasowski, P. Lacroute, J. Leng, J. Lian, H. Monahan, H. O'Geen, Z. Ouyang, E. C. Partridge, D. Patacsil, F. Pauli, D. Raha, L. Ramirez, T. E. Reddy, B. Reed, M. Shi, T. Slifer, J. Wang, L. Wu, X. Yang, K. Y. Yip, G. Zilberman-Schapira, S. Batzoglou, A. Sidow, P. J. Farnham,

- R. M. Myers, S. M. Weissman, and M. Snyder, 2012, Architecture of the human regulatory network derived from ENCODE data: *Nature*, v. 489, p. 91-100.
- Gey, G. O., W. D. Coffman, and M. T. Kubicek, 1952, Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium: *Cancer Res*, p. 264-265.
- Ghisletti, S., I. Barozzi, F. Mietton, S. Polletti, F. De Santa, E. Venturini, L. Gregory, L. Lonie, A. Chew, C. L. Wei, J. Ragoussis, and G. Natoli, 2010, Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages: *Immunity*, v. 32, p. 317-28.
- Granato, M., A. Farina, R. Gonnella, R. Santarelli, L. Frati, A. Faggioni, and A. Angeloni, 2006, Regulation of the expression of the Epstein-Barr virus early gene BFRF1: *Virology*, v. 347, p. 109-116.
- Greenspan, J. S., D. Greenspan, E. T. Lennette, D. I. Abrams, M. A. Conant, V. Petersen, and U. K. Freese, 1985, Replication of Epstein-Barr virus within the epithelial cells of oral "hairy" leukoplakia, an AIDS-associated lesion: *N Engl J Med*, v. 313, p. 1564-71.
- Grogan, E., H. Jenson, J. Countryman, L. Heston, L. Gradoville, and G. Miller, 1987, Transfection of a rearranged viral DNA fragment, WZhet, stably converts latent Epstein-Barr viral infection to productive infection in lymphoid cells: *Proceedings of the National Academy of Sciences of the United States of America*, v. 84, p. 1332-1336.
- Gruffat, H., F. Kadjouf, B. Mariamé, and E. Manet, 2012, The Epstein-Barr Virus BcRF1 Gene Product Is a TBP-Like Protein with an Essential Role in Late Gene Expression: *Journal of Virology*, v. 86, p. 6023-6032.
- Gruffat, H., E. Manet, and A. Sergeant, 2002, MEF2-mediated recruitment of class II HDAC at the EBV immediate early gene BZLF1 links latency and chromatin remodeling: *EMBO Rep*, v. 3, p. 141-6.
- Gruffat, H., R. Marchione, and E. Manet, 2016, Herpesvirus Late Gene Expression: A Viral-Specific Pre-initiation Complex Is Key: *Front Microbiol*, v. 7, p. 869.
- Guerra, C. A., P. W. Gikandi, A. J. Tatem, A. M. Noor, D. L. Smith, S. I. Hay, and R. W. Snow, 2008, The limits and intensity of *Plasmodium falciparum* transmission: Implications for malaria control and elimination worldwide: *PLoS Med*, v. 5, p. e38.
- Gulley, M. L., P. A. Eagan, L. Quintanilla-Martinez, A. L. Picado, B. N. Smir, C. Childs, C. D. Dunn, F. E. Craig, J. W. Williams, Jr., and P. M. Banks, 1994, Epstein-Barr virus DNA is abundant and monoclonal in the Reed-Sternberg cells of Hodgkin's disease: association with mixed cellularity subtype and Hispanic American ethnicity: *Blood*, v. 83, p. 1595-602.
- Gunnell, A., H. M. Webb, C. D. Wood, M. J. McClellan, B. Wichaidit, B. Kempkes, R. G. Jenner, C. Osborne, P. J. Farrell, and M. J. West, 2016, RUNX super-enhancer control through the Notch pathway by Epstein-Barr virus transcription factors regulates B cell growth: *Nucleic Acids Res*, v. 44, p. 4636-50.
- Gustems, M., A. Woellmer, U. Rothbauer, S. H. Eck, T. Wieland, D. Lutter, and W. Hammerschmidt, 2014, c-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs: *Nucleic Acids Research*, v. 42, p. 3059-3072.
- Hagemeier, S. R., S. J. Dickerson, Q. Meng, X. Yu, J. E. Mertz, and S. C. Kenney, 2010, Sumoylation of the Epstein-Barr virus BZLF1 protein inhibits its transcriptional activity and is regulated by the virus-encoded protein kinase: *J Virol*, v. 84, p. 4383-94.
- Hahn, A. M., L. E. Huye, S. Ning, J. Webster-Cyriaque, and J. S. Pagano, 2005, Interferon regulatory factor 7 is negatively regulated by the Epstein-Barr virus immediate-early gene, BZLF-1: *J Virol*, v. 79, p. 10040-52.

- Hammerschmidt, W., 2015, The Epigenetic Life Cycle of Epstein-Barr Virus, in C. Münz, ed., Epstein Barr Virus Volume 1: One Herpes Virus: Many Diseases: Cham, Springer International Publishing, p. 103-117.
- Hammerschmidt, W., and B. Sugden, 1988, Identification and characterization of oriLyt, a lytic origin of DNA replication of Epstein-Barr virus: *Cell*, v. 55, p. 427-33.
- Hammerschmidt, W., and B. Sugden, 2013, Replication of Epstein-Barr viral DNA: *Cold Spring Harb Perspect Biol*, v. 5, p. a013029.
- Haque, T., and D. H. Crawford, 1997, PCR amplification is more sensitive than tissue culture methods for Epstein-Barr virus detection in clinical material: *J Gen Virol*, v. 78 ( Pt 12), p. 3357-60.
- Hardwick, J. M., P. M. Lieberman, and S. D. Hayward, 1988, A new Epstein-Barr virus transactivator, R, induces expression of a cytoplasmic early antigen: *Journal of Virology*, v. 62, p. 2274-2284.
- Haring, M., S. Offermann, T. Danker, I. Horst, C. Peterhansel, and M. Stam, 2007, Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization: *Plant Methods*, v. 3, p. 11-11.
- Hatfull, G., A. T. Bankier, B. G. Barrell, and P. J. Farrell, 1988, Sequence analysis of Raji Epstein-Barr virus DNA: *Virology*, v. 164, p. 334-40.
- Heather, J., K. Flower, S. Isaac, and A. J. Sinclair, 2009, The Epstein-Barr virus lytic cycle activator Zta interacts with methylated ZRE in the promoter of host target gene *egr1*: *J Gen Virol*, v. 90, p. 1450-4.
- Heintzman, N. D., G. C. Hon, R. D. Hawkins, P. Kheradpour, A. Stark, L. F. Harp, Z. Ye, L. K. Lee, R. K. Stuart, C. W. Ching, K. A. Ching, J. E. Antosiewicz-Bourget, H. Liu, X. Zhang, R. D. Green, V. V. Lobanenko, R. Stewart, J. A. Thomson, G. E. Crawford, M. Kellis, and B. Ren, 2009, Histone modifications at human enhancers reflect global cell-type-specific gene expression: *Nature*, v. 459, p. 108-12.
- Heller, M., E. Flemington, E. Kieff, and P. Deininger, 1985, Repeat arrays in cellular DNA related to the Epstein-Barr virus IR3 repeat: *Mol Cell Biol*, v. 5, p. 457-65.
- Heller, M., A. Henderson, and E. Kieff, 1982, Repeat array in Epstein-Barr virus DNA is related to cell DNA sequences interspersed on human chromosomes: *Proceedings of the National Academy of Sciences of the United States of America*, v. 79, p. 5916-5920.
- Henderson, S., D. Huen, M. Rowe, C. Dawson, G. Johnson, and A. Rickinson, 1993, Epstein-Barr virus-coded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death: *Proceedings of the National Academy of Sciences of the United States of America*, v. 90, p. 8479-8483.
- Henle, G., and W. Henle, 1966, Immunofluorescence in Cells Derived from Burkitt's Lymphoma: *Journal of Bacteriology*, v. 91, p. 1248-1256.
- Henle, G., W. Henle, and V. Diehl, 1968, Relation of Burkitt's tumor-associated herpes-type virus to infectious mononucleosis: *Proc Natl Acad Sci U S A*, v. 59, p. 94-101.
- Henle, W., V. Diehl, G. Kohn, H. Zur Hausen, and G. Henle, 1967, Herpes-type virus and chromosome marker in normal leukocytes after growth with irradiated Burkitt cells: *Science*, v. 157, p. 1064-5.
- Henle, W., G. Henle, H. C. Ho, P. Burtin, Y. Cachin, P. Clifford, A. de Schryver, G. de The, V. Diehl, and G. Klein, 1970, Antibodies to Epstein-Barr virus in nasopharyngeal carcinoma, other head and neck neoplasms, and control groups: *J Natl Cancer Inst*, v. 44, p. 225-31.
- Herz, H.-M., D. Hu, and A. Shilatifard, 2014, Enhancer Malfunction in Cancer: *Molecular cell*, v. 53, p. 859-866.
- Heston, L., A. El-Guindy, J. Countryman, C. Dela Cruz, H.-J. Delecluse, and G. Miller, 2006, Amino acids in the basic domain of Epstein-Barr virus ZEBRA protein play distinct roles in DNA binding, activation of early lytic gene expression, and promotion of viral DNA replication: *Journal of Virology*, v. 80, p. 9115-9133.



- Hickey, S. M., and V. C. Strasburger, 1997, What every paediatrician should know about infectious mononucleosis in adolescents: *Pediatric Clinics of North America*, v. 44, p. 1541-1556.
- Hislop, A. D., M. E. Rensing, D. van Leeuwen, V. A. Pudney, D. Horst, D. Koppers-Lalic, N. P. Croft, J. J. Neefjes, A. B. Rickinson, and E. J. Wiertz, 2007, A CD8+ T cell immune evasion protein specific to Epstein-Barr virus and its close relatives in Old World primates: *J Exp Med*, v. 204, p. 1863-73.
- Hnisz, D., Brian J. Abraham, Tong I. Lee, A. Lau, V. Saint-André, Alla A. Sigova, Heather A. Hoke, and Richard A. Young, 2013, Super-Enhancers in the Control of Cell Identity and Disease: *Cell*, v. 155, p. 934-947.
- Hodgkin, 1832, On some Morbid Appearances of the Absorbent Glands and Spleen: *Medico-chirurgical transactions*, v. 17, p. 68-114.
- Holley-Guthrie, E. A., E. B. Quinlivan, E. C. Mar, and S. Kenney, 1990, The Epstein-Barr virus (EBV) BMRF1 promoter for early antigen (EA-D) is regulated by the EBV transactivators, BRLF1 and BZLF1, in a cell-specific manner: *J Virol*, v. 64, p. 3753-9.
- Homouz, D., and A. S. Kudlicki, 2013, The 3D organization of the yeast genome correlates with co-expression and reflects functional relations between genes: *PLoS One*, v. 8, p. e54699.
- Hong, G. K., H. J. Delecluse, H. Gruffat, T. E. Morrison, W. H. Feng, A. Sergeant, and S. C. Kenney, 2004, The BRRF1 early gene of Epstein-Barr virus encodes a transcription factor that enhances induction of lytic infection by BRLF1: *J Virol*, v. 78, p. 4983-92.
- Hopwood, P., and D. Crawford, 2000, The role of EBV in post-transplant malignancies: a review: *Journal of Clinical Pathology*, v. 53, p. 248-254.
- Horst, D., D. van Leeuwen, N. P. Croft, M. A. Garstka, A. D. Hislop, E. Kremmer, A. B. Rickinson, E. J. Wiertz, and M. E. Rensing, 2009, Specific targeting of the EBV lytic phase protein BNLF2a to the transporter associated with antigen processing results in impairment of HLA class I-restricted antigen presentation: *J Immunol*, v. 182, p. 2313-24.
- Hsu, M., S. Y. Wu, S. S. Chang, I. J. Su, C. H. Tsai, S. J. Lai, A. L. Shiau, K. Takada, and Y. Chang, 2008, Epstein-Barr virus lytic transactivator Zta enhances chemotactic activity through induction of interleukin-8 in nasopharyngeal carcinoma cells: *J Virol*, v. 82, p. 3679-88.
- Huang, Q., C. Gong, J. Li, Z. Zhuo, Y. Chen, J. Wang, and Z.-C. Hua, 2012, Distance and Helical Phase Dependence of Synergistic Transcription Activation in cis-Regulatory Module: *PLoS ONE*, v. 7, p. e31198.
- Hudson, G. S., P. J. Farrell, and B. G. Barrell, 1985, Two related but differentially expressed potential membrane proteins encoded by the EcoRI Dhet region of Epstein-Barr virus B95-8: *Journal of Virology*, v. 53, p. 528-535.
- Hummel, M., S. Bentink, H. Berger, W. Klapper, S. Wessendorf, T. F. E. Barth, H.-W. Bernd, S. B. Cogliatti, J. Dierlamm, A. C. Feller, M.-L. Hansmann, E. Haralambieva, L. Harder, D. Hasenclever, M. Kühn, D. Lenze, P. Lichter, J. I. Martin-Subero, P. Möller, H.-K. Müller-Hermelink, G. Ott, R. M. Parwaresch, C. Pott, A. Rosenwald, M. Rosolowski, C. Schwaenen, B. Stürzenhake, M. Szczepanowski, H. Trautmann, H.-H. Wacker, R. Spang, M. Loeffler, L. Trümper, H. Stein, and R. Siebert, 2006, A Biologic Definition of Burkitt's Lymphoma from Transcriptional and Genomic Profiling: *New England Journal of Medicine*, v. 354, p. 2419-2430.
- Hung, C. H., and S. T. Liu, 1999, Characterization of the Epstein-Barr virus BALF2 promoter: *J Gen Virol*, v. 80 ( Pt 10), p. 2747-50.
- Huo, Q., N. Zhang, and Q. Yang, 2012, Epstein-Barr Virus Infection and Sporadic Breast Cancer Risk: A Meta-Analysis: *PLoS ONE*, v. 7, p. e31656.
- Hurley, E. A., and D. A. Thorley-Lawson, 1988, B cell activation and the establishment of Epstein-Barr virus latency: *J Exp Med*, v. 168, p. 2059-75.

- Iizasa, H., A. Nanbo, J. Nishikawa, M. Jinushi, and H. Yoshiyama, 2012, Epstein-Barr Virus (EBV)-associated Gastric Carcinoma: *Viruses*, v. 4, p. 3420-3439.
- Imai, S., 1994, Gastric carcinoma: monoclonal epithelial malignant cells expressing Epstein-Barr virus latent infection protein: *Proc. Natl Acad. Sci. USA*, v. 91, p. 9131-9135.
- Inman, G. J., U. K. Binne, G. A. Parker, P. J. Farrell, and M. J. Allday, 2001, Activators of the Epstein-Barr virus lytic program concomitantly induce apoptosis, but lytic gene expression protects from cell death: *J Virol*, v. 75, p. 2400-10.
- Inoue, F., and N. Ahituv, 2015, Decoding enhancers using massively parallel reporter assays: *Genomics*, v. 106, p. 159-164.
- Janson, L., and U. Pettersson, 1990, Cooperative interactions between transcription factors Sp1 and OTF-1: *Proceedings of the National Academy of Sciences of the United States of America*, v. 87, p. 4732-4736.
- Jenkins, T. D., H. Nakagawa, and A. K. Rustgi, 1997, The keratinocyte-specific Epstein-Barr virus ED-L2 promoter is regulated by phorbol 12-myristate 13-acetate through two cis-regulatory elements containing E-box and Kruppel-like factor motifs: *J Biol Chem*, v. 272, p. 24433-42.
- Jenkins, T. D., O. G. Opitz, J. Okano, and A. K. Rustgi, 1998, Transactivation of the human keratin 4 and Epstein-Barr virus ED-L2 promoters by gut-enriched Kruppel-like factor: *J Biol Chem*, v. 273, p. 10747-54.
- Jiang, R., R. S. Scott, and L. M. Hutt-Fletcher, 2006, Epstein-Barr virus shed in saliva is high in B-cell-tropic glycoprotein gp42: *J Virol*, v. 80, p. 7281-3.
- Jochum, S., A. Moosmann, S. Lang, W. Hammerschmidt, and R. Zeidler, 2012, The EBV immunoevasins vIL-10 and BNLF2a protect newly infected B cells from immune recognition and elimination: *PLoS Pathog*, v. 8, p. e1002704.
- Jones, J. F., S. Shurin, C. Abramowsky, R. R. Tubbs, C. G. Sciotto, R. Wahl, J. Sands, D. Gottman, B. Z. Katz, and J. Sklar, 1988, T-Cell Lymphomas Containing Epstein-Barr Viral DNA in Patients with Chronic Epstein-Barr Virus Infections: *New England Journal of Medicine*, v. 318, p. 733-741.
- Jung, Y. J., H. Choi, H. Kim, and S. K. Lee, 2014, MicroRNA miR-BART20-5p stabilizes Epstein-Barr virus latency by directly targeting BZLF1 and BRLF1: *J Virol*, v. 88, p. 9027-37.
- Kaczynski, J., T. Cook, and R. Urrutia, 2003, Sp1- and Kruppel-like transcription factors: *Genome Biol*, v. 4, p. 206.
- Kalla, M., C. Gobel, and W. Hammerschmidt, 2012, The lytic phase of Epstein-Barr virus requires a viral genome with 5-methylcytosine residues in CpG sites: *J Virol*, v. 86, p. 447-58.
- Kalla, M., A. Schmeinck, M. Bergbauer, D. Pich, and W. Hammerschmidt, 2010, AP-1 homolog BZLF1 of Epstein-Barr virus has two essential functions dependent on the epigenetic state of the viral genome: *Proc Natl Acad Sci U S A*, v. 107, p. 850-5.
- Kang, M.-S., and E. Kieff, 2015, Epstein-Barr virus latent genes: *Exp Mol Med*, v. 47, p. e131.
- Kapatai, G., and P. Murray, 2007, Contribution of the Epstein-Barr virus to the molecular pathogenesis of Hodgkin lymphoma: *Journal of Clinical Pathology*, v. 60, p. 1342-1349.
- Karimi, L., D. H. Crawford, S. Speck, and L. J. Nicholson, 1995, Identification of an epithelial cell differentiation responsive region within the BZLF1 promoter of the Epstein-Barr virus: *J Gen Virol*, v. 76 ( Pt 4), p. 759-65.
- Karlsson, Q. H., C. Schelcher, E. Verrall, C. Petosa, and A. J. Sinclair, 2008a, Methylated DNA Recognition during the Reversal of Epigenetic Silencing Is Regulated by Cysteine and Serine Residues in the Epstein-Barr Virus Lytic Switch Protein: *PLoS Pathog*, v. 4.

- Karlsson, Q. H., C. Schelcher, E. Verrall, C. Petosa, and A. J. Sinclair, 2008b, The reversal of epigenetic silencing of the EBV genome is regulated by viral bZIP protein: *Biochem Soc Trans*, v. 36, p. 637-9.
- Kaufman, H. L., F. J. Kohlhapp, and A. Zloza, 2015, Oncolytic viruses: a new class of immunotherapy drugs: *Nat Rev Drug Discov*, v. 14, p. 642-62.
- Kelly, G. L., A. E. Milner, R. J. Tierney, D. S. Croom-Carter, M. Altmann, W. Hammerschmidt, A. I. Bell, and A. B. Rickinson, 2005, Epstein-Barr virus nuclear antigen 2 (EBNA2) gene deletion is consistently linked with EBNA3A, -3B, and -3C expression in Burkitt's lymphoma cells and with increased resistance to apoptosis: *J Virol*, v. 79, p. 10709-17.
- Kempkes, B., and P. D. Ling, 2015, EBNA2 and Its Coactivator EBNA-LP, *in* C. Münz, ed., *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 35-59.
- Kenney, S., J. Kamine, E. Holley-Guthrie, J. C. Lin, E. C. Mar, and J. Pagano, 1989, The Epstein-Barr virus (EBV) BZLF1 immediate-early gene product differentially affects latent versus productive EBV promoters: *J Virol*, v. 63, p. 1729-36.
- Kenney, S. C., 2007, Reactivation and lytic replication of EBV: *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*.
- Kenney, S. C., E. Holley-Guthrie, E. B. Quinlivan, D. Gutsch, Q. Zhang, T. Bender, J. F. Giot, and A. Sergeant, 1992, The cellular oncogene c-myc can interact synergistically with the Epstein-Barr virus BZLF1 transactivator in lymphoid cells: *Mol Cell Biol*, v. 12, p. 136-46.
- Kenney, S. C., and J. E. Mertz, 2014, Regulation of the latent-lytic switch in Epstein-Barr virus: *Seminars in cancer biology*, v. 0, p. 60-68.
- Kheradpour, P., J. Ernst, A. Melnikov, P. Rogov, L. Wang, X. Zhang, J. Alston, T. S. Mikkelsen, and M. Kellis, 2013, Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay: *Genome Research*, v. 23, p. 800-811.
- Kidder, B. L., G. Hu, and K. Zhao, 2011, ChIP-Seq: Technical Considerations for Obtaining High Quality Data: *Nature immunology*, v. 12, p. 918-922.
- Kieser, A., and K. R. Sterz, 2015, The Latent Membrane Protein 1 (LMP1): *Curr Top Microbiol Immunol*, v. 391, p. 119-49.
- Kilger, E., A. Kieser, M. Baumann, and W. Hammerschmidt, 1998, Epstein-Barr virus-mediated B-cell proliferation is dependent upon latent membrane protein 1, which simulates an activated CD40 receptor: *EMBO J.*, v. 17, p. 1700-1709.
- Klein, G., L. Dombos, and B. Gothoskar, 1972, Sensitivity of Epstein-Barr virus (EBV)producer and non-producer human lymphoblastoid cell lines to superinfection with EB-virus: *Int J Cancer*, v. 10.
- Klug, M., and M. Rehli, 2006, Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector: *Epigenetics*, v. 1, p. 127-30.
- Komano, J., S. Maruo, K. Kurozumi, T. Oda, and K. Takada, 1999, Oncogenic role of Epstein-Barr virus-encoded RNAs in Burkitt's lymphoma cell line Akata: *J Virol*, v. 73, p. 9827-9831.
- Kouzarides, T., G. Packham, A. Cook, and P. J. Farrell, 1991, The BZLF1 protein of EBV has a coiled coil dimerisation domain without a heptad leucine repeat but with homology to the C/EBP leucine zipper: *Oncogene*, v. 6, p. 195-204.
- Kraus, R. J., J. G. Perrigoue, and J. E. Mertz, 2003, ZEB negatively regulates the lytic-switch BZLF1 gene promoter of Epstein-Barr virus: *Journal of Virology*, v. 77, p. 199-207.
- Krivega, I., and A. Dean, 2012, Enhancer and promoter interactions — long distance calls: *Current opinion in genetics & development*, v. 22, p. 79-85.
- Kudoh, A., M. Fujita, T. Kiyono, K. Kuzushima, Y. Sugaya, S. Izuta, Y. Nishiyama, and T. Tsurumi, 2003, Reactivation of lytic replication from B cells latently infected with Epstein-Barr virus occurs with high S-phase cyclin-dependent kinase

- activity while inhibiting cellular DNA replication: *Journal of Virology*, v. 77, p. 851-861.
- Kuppers, R., 2012, New insights in the biology of Hodgkin lymphoma: *Hematology Am Soc Hematol Educ Program*, v. 2012, p. 328-34.
- Kuppers, R., A. Engert, and M. L. Hansmann, 2012, Hodgkin lymphoma: *J Clin Invest*, v. 122, p. 3439-47.
- Kurth, J., T. Spieker, J. Wustrow, J. G. Strickler, M.-L. Hansmann, K. Rajewsky, and R. Küppers, 2000, EBV-Infected B Cells in Infectious Mononucleosis: Viral Strategies for Spreading in the B Cell Compartment and Establishing Latency: *Immunity*, v. 13, p. 485-495.
- Kutok, J. L., and F. Wang, 2006, Spectrum of Epstein-Barr virus-associated diseases: *Annu Rev Pathol*, v. 1, p. 375-404.
- Kuwana, Y., M. Takei, M. Yajima, K.-I. Imadome, H. Inomata, M. Shiozaki, N. Ikumi, T. Nozaki, H. Shiraiwa, N. Kitamura, J. Takeuchi, S. Sawada, N. Yamamoto, N. Shimizu, M. Ito, and S. Fujiwara, 2011, Epstein-Barr Virus Induces Erosive Arthritis in Humanized Mice: *PLoS ONE*, v. 6, p. e26630.
- Kuzembayeva, M., M. Hayes, and B. Sugden, 2014, Multiple functions are mediated by the miRNAs of Epstein-Barr virus: *Curr Opin Virol*, v. 7, p. 61-5.
- Kwok, H., A. H. Tong, C. H. Lin, S. Lok, P. J. Farrell, D. L. Kwong, and A. K. Chiang, 2012, Genomic sequencing and comparative analysis of Epstein-Barr virus genome isolated from primary nasopharyngeal carcinoma biopsy: *PLoS One*, v. 7, p. e36939.
- Lagenaur, L. A., and J. M. Palefsky, 1999, Regulation of Epstein-Barr virus promoters in oral epithelial cells and lymphocytes: *J Virol*, v. 73, p. 6566-72.
- Laichalk, L. L., and D. A. Thorley-Lawson, 2005, Terminal differentiation into plasma cells initiates the replicative cycle of Epstein-Barr virus in vivo: *J Virol*, v. 79, p. 1296-1307.
- Landt, S. G., G. K. Marinov, A. Kundaje, P. Kheradpour, F. Pauli, S. Batzoglou, B. E. Bernstein, P. Bickel, J. B. Brown, P. Cayting, Y. Chen, G. DeSalvo, C. Epstein, K. I. Fisher-Aylor, G. Euskirchen, M. Gerstein, J. Gertz, A. J. Hartemink, M. M. Hoffman, V. R. Iyer, Y. L. Jung, S. Karmakar, M. Kellis, P. V. Kharchenko, Q. Li, T. Liu, X. S. Liu, L. Ma, A. Milosavljevic, R. M. Myers, P. J. Park, M. J. Pazin, M. D. Perry, D. Raha, T. E. Reddy, J. Rozowsky, N. Shores, A. Sidow, M. Slattery, J. A. Stamatoyannopoulos, M. Y. Tolstorukov, K. P. White, S. Xi, P. J. Farnham, J. D. Lieb, B. J. Wold, and M. Snyder, 2012, ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia: *Genome Research*, v. 22, p. 1813-1831.
- Lau, R., J. Middeldorp, and P. J. Farrell, 1993, Epstein-Barr virus gene expression in oral hairy leukoplakia: *Virology*, v. 195, p. 463-74.
- Lee, T. I., and R. A. Young, 2000, Transcription of eukaryotic protein-coding genes: *Annu Rev Genet*, v. 34, p. 77-137.
- Lelli, K. M., M. Slattery, and R. S. Mann, 2012, Disentangling the many layers of eukaryotic transcriptional regulation: *Annu Rev Genet*, v. 46, p. 43-68.
- Lemon, S. M., L. M. Hutt, J. E. Shaw, J. L. Li, and J. S. Pagano, 1977, Replication of EBV in epithelial cells during infectious mononucleosis: *Nature*, v. 268, p. 268-70.
- Lemon, S. M., L. M. Hutt, J. E. Shaw, J. L. Li, and J. S. Pagano, 1978, Replication of Epstein-Barr virus DNA in epithelial cells in vivo: *IARC Sci Publ*, p. 739-44.
- Lerner, M. R., N. C. Andrews, G. Miller, and J. A. Steitz, 1981, Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus: *Proc Natl Acad Sci U S A*, v. 78, p. 805-9.
- Levitskaya, J., M. Coram, V. Levitsky, S. Imreh, P. M. Steigerwald-Mullen, G. Klein, M. G. Kurilla, and M. G. Masucci, 1995, Inhibition of antigen processing by the



- internal repeat region of the Epstein-Barr virus nuclear antigen-1: *Nature*, v. 375, p. 685-8.
- Levitskaya, J., A. Sharipo, A. Leonchiks, A. Ciechanover, and M. G. Masucci, 1997, Inhibition of ubiquitin/proteasome-dependent protein degradation by the Gly-Ala repeat domain of the Epstein-Barr virus nuclear antigen 1: *Proc Natl Acad Sci USA*, v. 94, p. 12616-12621.
- Li, D., L. Qian, C. Chen, M. Shi, M. Yu, M. Hu, L. Song, B. Shen, and N. Guo, 2009, Down-regulation of MHC class II expression through inhibition of CIITA transcription by lytic transactivator Zta during Epstein-Barr virus reactivation: *J Immunol*, v. 182, p. 1799-809.
- Li, Q. X., L. S. Young, G. Niedobitek, C. W. Dawson, M. Birkenbach, F. Wang, and A. B. Rickinson, 1992, Epstein-Barr virus infection and replication in a human epithelial cell system: *Nature*, v. 356, p. 347-50.
- Li, R.-C., Y. Du, Q.-Y. Zeng, L.-Q. Tang, H. Zhang, Y. Li, W.-L. Liu, Q. Zhong, M.-S. Zeng, and X.-M. Huang, 2016, Epstein-Barr virus glycoprotein gH/gL antibodies complement IgA-viral capsid antigen for diagnosis of nasopharyngeal carcinoma: *Oncotarget*, v. 7, p. 16372-16383.
- Liao, G., J. Huang, E. D. Fixman, and S. D. Hayward, 2005, The Epstein-Barr virus replication protein BBLF2/3 provides an origin-tethering function through interaction with the zinc finger DNA binding protein ZBRK1 and the KAP-1 corepressor: *J Virol*, v. 79, p. 245-56.
- Liao, G., F. Y. Wu, and S. D. Hayward, 2001, Interaction with the Epstein-Barr virus helicase targets Zta to DNA replication compartments: *J Virol*, v. 75, p. 8792-802.
- Lieberman, P., 1994, Identification of functional targets of the Zta transcriptional activator by formation of stable preinitiation complex intermediates: *Mol Cell Biol*, v. 14, p. 8365-75.
- Lieberman, P. M., 2013, Keeping it quiet: chromatin control of gammaherpesvirus latency: *Nat Rev Microbiol*, v. 11, p. 863-75.
- Lieberman, P. M., 2015, Chromatin Structure of Epstein-Barr Virus Latent Episomes, *in* C. Münz, ed., *Epstein Barr Virus Volume 1: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 71-102.
- Lieberman, P. M., and A. J. Berk, 1990, In vitro transcriptional activation, dimerization, and DNA-binding specificity of the Epstein-Barr virus Zta protein: *J Virol*, v. 64, p. 2560-8.
- Lieberman, P. M., and A. J. Berk, 1991, The Zta trans-activator protein stabilizes TFIID association with promoter DNA by direct protein-protein interaction: *Genes Dev*, v. 5, p. 2441-54.
- Lieberman, P. M., and A. J. Berk, 1994, A mechanism for TAFs in transcriptional activation: activation domain enhancement of TFIID-TFIIA--promoter DNA complex formation: *Genes Dev*, v. 8, p. 995-1006.
- Lieberman, P. M., J. M. Hardwick, and S. D. Hayward, 1989, Responsiveness of the Epstein-Barr virus NotI repeat promoter to the Z transactivator is mediated in a cell-type-specific manner by two independent signal regions: *Journal of Virology*, v. 63, p. 3040-3050.
- Lieberman, P. M., J. M. Hardwick, J. Sample, G. S. Hayward, and S. D. Hayward, 1990, The zta transactivator involved in induction of lytic cycle gene expression in Epstein-Barr virus-infected lymphocytes binds to both AP-1 and ZRE sites in target promoter and enhancer regions: *J Virol*, v. 64, p. 1143-55.
- Lieberman, P. M., J. Ozer, and D. B. Gursel, 1997, Requirement for transcription factor IIA (TFIIA)-TFIID recruitment by an activator depends on promoter structure and template competition: *Mol Cell Biol*, v. 17, p. 6624-32.
- Lin, S.-F., T.-Y. Hsu, M.-Y. Liu, L.-S. Lin, H.-L. Yang, J.-Y. Chen, and C.-S. Yang, 1995, Characterization of Epstein-Barr Virus DNase and Its Interaction with the Major DNA Binding Protein: *Virology*, v. 208, p. 712-722.

- Lin, Z., and E. K. Flemington, 2010, Regulation of EBV latency by viral lytic proteins: Epstein-Barr virus: latency and transformation. Caister Academic Press, Norwich, United Kingdom, p. 167-192.
- Liu, P., and S. H. Speck, 2003, Synergistic autoactivation of the Epstein-Barr virus immediate-early BRLF1 promoter by Rta and Zta: *Virology*, v. 310, p. 199-206.
- Longnecker, R. M., E. Kieff, and J. I. Cohen, 2013, Epstein-barr virus, *Fields Virology: Sixth Edition*, v. 1.
- Lopez, G., F. Schaufele, P. Webb, J. M. Holloway, J. D. Baxter, and P. J. Kushner, 1993, Positive and negative modulation of Jun action by thyroid hormone receptor at a unique AP1 site: *Molecular and Cellular Biology*, v. 13, p. 3042-3049.
- Lorenzetti, M. A., M. Gantuz, J. Altcheh, E. De Matteo, P. A. Chabay, and M. V. Preciado, 2014, Epstein-Barr virus BZLF1 gene polymorphisms: malignancy related or geographically distributed variants?: *Clinical Microbiology and Infection*, v. 20, p. O861-O869.
- Ludwig, M. Z., N. H. Patel, and M. Kreitman, 1998, Functional analysis of eve stripe 2 enhancer evolution in *Drosophila*: rules governing conservation and change: *Development*, v. 125, p. 949-58.
- Lunemann, J. D., T. Kamradt, R. Martin, and C. Munz, 2007, Epstein-barr virus: environmental trigger of multiple sclerosis?: *J Virol*, v. 81, p. 6777-84.
- Luzuriaga, K., and J. L. Sullivan, 2010, Infectious Mononucleosis: *New England Journal of Medicine*, v. 362, p. 1993-2000.
- Löning, T., R.-P. Henke, P. Reichart, and J. Becker, 1987, In situ hybridization to detect Epstein-Barr virus DNA in oral tissues of HIV-infected patients: *Virchows Archiv A*, v. 412, p. 127-133.
- Mackenzie, I. S., S. V. Morant, G. A. Bloomfield, T. M. MacDonald, and J. O'Riordan, 2014, Incidence and prevalence of multiple sclerosis in the UK 1990-2010: a descriptive study in the General Practice Research Database: *J Neurol Neurosurg Psychiatry*, v. 85, p. 76-84.
- Magrath, I., 2012, Epidemiology: clues to the pathogenesis of Burkitt lymphoma: *Br J Haematol*, v. 156, p. 744-56.
- Mahot, S., A. Sergeant, E. Drouet, and H. Gruffat, 2003a, A novel function for the Epstein-Barr virus transcription factor EB1/Zta: Induction of transcription of the hIL-10 gene: *Journal of General Virology*, v. 84, p. 965-974.
- Mahot, S., A. Sergeant, E. Drouet, and H. Gruffat, 2003b, A novel function for the Epstein-Barr virus transcription factor EB1/Zta: induction of transcription of the hIL-10 gene: *J Gen Virol*, v. 84, p. 965-74.
- Marrão, G., M. Habib, A. Paiva, D. Bicout, C. Fallecker, S. Franco, S. Fafi-Kremer, T. Simões da Silva, P. Morand, C. Freire de Oliveira, and E. Drouet, 2014, Epstein-Barr virus infection and clinical outcome in breast cancer patients correlate with immune cell TNF- $\alpha$ /IFN- $\gamma$  response: *BMC Cancer*, v. 14, p. 1-11.
- Marschall, M., U. Leser, R. Seibl, and H. Wolf, 1989, Identification of proteins encoded by Epstein-Barr virus trans-activator genes: *J Virol*, v. 63, p. 938-42.
- Martini, M., D. Capello, D. Serraino, A. Navarra, F. Pierconti, T. Cenci, G. Gaidano, and L. M. Larocca, 2007, Characterization of variants in the promoter of EBV gene BZLF1 in normal donors, HIV-positive patients and in AIDS-related lymphomas: *Journal of Infection*, v. 54, p. 298-306.
- Massini, G., D. Siemer, and S. Hohaus, 2009, EBV in Hodgkin Lymphoma: *Mediterranean Journal of Hematology and Infectious Diseases*, v. 1, p. e2009013.
- Maston, G. A., S. K. Evans, and M. R. Green, 2006, Transcriptional regulatory elements in the human genome: *Annu Rev Genomics Hum Genet*, v. 7, p. 29-59.
- Matsuura, H., A. N. Kirschner, R. Longnecker, and T. S. Jardetzky, 2010, Crystal structure of the Epstein-Barr virus (EBV) glycoprotein H/glycoprotein L (gH/gL)

- complex: Proceedings of the National Academy of Sciences of the United States of America, v. 107, p. 22641-22646.
- Mauser, A., S. i. Saito, E. Appella, C. W. Anderson, W. T. Seaman, and S. Kenney, 2002, The Epstein-Barr Virus Immediate-Early Protein BZLF1 Regulates p53 Function through Multiple Mechanisms: *Journal of Virology*, v. 76, p. 12503-12512.
- McClellan, M. J., C. D. Wood, O. Ojieniyi, T. J. Cooper, A. Kanhere, A. Arvey, H. M. Webb, R. D. Palermo, M. L. Harth-Hertle, B. Kempkes, R. G. Jenner, and M. J. West, 2013, Modulation of enhancer looping and differential gene targeting by Epstein-Barr virus transcription factors directs cellular reprogramming: *PLoS Pathog*, v. 9, p. e1003636.
- McDonald, C., C. E. Karstegl, P. Kellam, and P. J. Farrell, 2010, Regulation of the Epstein-Barr virus Zp promoter in B lymphocytes during reactivation from latency: *J Gen Virol*, v. 91, p. 622-9.
- McDonald, C. M., C. Petosa, and P. J. Farrell, 2009, Interaction of Epstein-Barr virus BZLF1 C-terminal tail structure and core zipper is required for DNA replication but not for promoter transactivation: *Journal of Virology*, v. 83, p. 3397-3401.
- McGeoch, D. J., S. Cook, A. Dolan, F. E. Jamieson, and E. A. R. Telford, 1995, Molecular Phylogeny and Evolutionary Timescale for the Family of Mammalian Herpesviruses: *Journal of Molecular Biology*, v. 247, p. 443-458.
- McKenzie, J., and A. El-Guindy, 2015, Epstein-Barr Virus Lytic Cycle Reactivation, *in* C. Münz, ed., *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 237-261.
- Melnikov, A., A. Murugan, X. Zhang, T. Tesileanu, L. Wang, P. Rogov, S. Feizi, A. Gnirke, C. G. Callan, Jr., J. B. Kinney, M. Kellis, E. S. Lander, and T. S. Mikkelsen, 2012, Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay: *Nat Biotechnol*, v. 30, p. 271-7.
- Meng, Q., S. R. Hagemeier, J. D. Fingerioth, E. Gershburg, J. S. Pagano, and S. C. Kenney, 2010, The Epstein-Barr virus (EBV)-encoded protein kinase, EBV-PK, but not the thymidine kinase (EBV-TK), is required for ganciclovir and acyclovir inhibition of lytic viral production: *Journal of Virology*, v. 84, p. 4534-4542.
- Metzenberg, S., 1989, Relative rates of RNA synthesis across the genome of Epstein-Barr virus are highest near oriP and oriLyt: *J Virol*, v. 63, p. 4938-44.
- Miller, G., A. El-Guindy, J. Countryman, J. Ye, and L. Gradoville, 2007, Lytic cycle switches of oncogenic human gammaherpesviruses: *Adv Cancer Res*, v. 97, p. 81-109.
- Montone, K. T., R. L. Hodinka, K. E. Salhany, E. Lavi, A. Rostami, and J. E. Tomaszewski, 1996, Identification of Epstein-Barr virus lytic activity in post-transplantation lymphoproliferative disease: *Mod Pathol*, v. 9, p. 621-30.
- Moon, U. Y., S. J. Park, S. T. Oh, W.-U. Kim, S.-H. Park, S.-H. Lee, C.-S. Cho, H.-Y. Kim, W.-K. Lee, and S. K. Lee, 2004, Patients with systemic lupus erythematosus have abnormally elevated Epstein-Barr virus load in blood: *Arthritis Research & Therapy*, v. 6, p. R295-R302.
- Moore, M. D., R. G. DiScipio, N. R. Cooper, and G. R. Nemerow, 1989, Hydrodynamic, electron microscopic, and ligand-binding analysis of the Epstein-Barr virus/C3dg receptor (CR2): *J Biol Chem*, v. 264, p. 20576-82.
- Moore, P. S., and Y. Chang, 2010, Why do viruses cause cancer? Highlights of the first century of human tumour virology: *Nature reviews. Cancer*, v. 10, p. 878-889.
- Morrison, T. E., and S. C. Kenney, 2004, BZLF1, an Epstein-Barr virus immediate-early protein, induces p65 nuclear translocation while inhibiting p65 transcriptional function: *Virology*, v. 328, p. 219-32.
- Murata, T., Y. Kondo, A. Sugimoto, D. Kawashima, S. Saito, H. Isomura, T. Kanda, and T. Tsurumi, 2012, Epigenetic Histone Modification of Epstein-Barr Virus BZLF1

- Promoter during Latency and Reactivation in Raji Cells: *Journal of Virology*, v. 86, p. 4752-4761.
- Murata, T., Y. Narita, A. Sugimoto, D. Kawashima, T. Kanda, and T. Tsurumi, 2013, Contribution of myocyte enhancer factor 2 family transcription factors to BZLF1 expression in Epstein-Barr virus reactivation from latency: *J Virol*, v. 87, p. 10148-62.
- Murata, T., and T. Tsurumi, 2014, Switching of EBV cycles between latent and lytic states: *Rev Med Virol*, v. 24, p. 142-53.
- Murray, P., and A. Bell, 2015, Contribution of the Epstein-Barr Virus to the Pathogenesis of Hodgkin Lymphoma: *Curr Top Microbiol Immunol*, v. 390, p. 287-313.
- Najand, N., J.-R. Ryu, and W. J. Brook, 2012, In Vitro Site Selection of a Consensus Binding Site for the *Drosophila melanogaster* Tbx20 Homolog Midline: *PLoS ONE*, v. 7, p. e48176.
- Nakagawa, H., T. Inomoto, and A. K. Rustgi, 1997a, A CACCC Box-like cis-Regulatory Element of the Epstein-Barr Virus ED-L2 Promoter Interacts with a Novel Transcriptional Factor in Tissue-specific Squamous Epithelia: *Journal of Biological Chemistry*, v. 272, p. 16688-16699.
- Nakagawa, H., T. C. Wang, L. Zukerberg, R. Odze, K. Togawa, G. H. May, J. Wilson, and A. K. Rustgi, 1997b, The targeting of the cyclin D1 oncogene by an Epstein-Barr virus promoter in transgenic mice causes dysplasia in the tongue, esophagus and forestomach: *Oncogene*, v. 14, p. 1185-90.
- Nawandar, D. M., A. Wang, K. Makielski, D. Lee, S. Ma, E. Barlow, J. Reusch, R. Jiang, C. K. Wille, D. Greenspan, J. S. Greenspan, J. E. Mertz, L. Hutt-Fletcher, E. C. Johannsen, P. F. Lambert, and S. C. Kenney, 2015, Differentiation-Dependent KLF4 Expression Promotes Lytic Epstein-Barr Virus Infection in Epithelial Cells: *PLoS Pathog*, v. 11, p. e1005195.
- Neefjes, J., M. L. Jongsma, P. Paul, and O. Bakke, 2011, Towards a systems understanding of MHC class I and MHC class II antigen presentation: *Nat Rev Immunol*, v. 11, p. 823-36.
- Nemerow, G. R., C. Mold, V. K. Schwend, V. Tollefson, and N. R. Cooper, 1987, Identification of gp350 as the viral glycoprotein mediating attachment of Epstein-Barr virus (EBV) to the EBV/C3d receptor of B cells: Sequence homology of gp350 and C3 complement fragment C3d: *Journal of Virology*, v. 61, p. 1416-1420.
- Neri, A., F. Barriga, G. Inghirami, D. M. Knowles, J. Neequaye, I. T. Magrath, and R. Dalla-Favera, 1991, Epstein-Barr virus infection precedes clonal expansion in Burkitt's and acquired immunodeficiency syndrome-associated lymphoma: *Blood*, v. 77, p. 1092-5.
- Nguyen, V. T., M. Morange, and O. Bensaude, 1988, Firefly luciferase luminescence assays using scintillation counters for quantitation in transfected mammalian cells: *Anal Biochem*, v. 171, p. 404-8.
- Niedobitek, G., A. Agathangelou, M. Rowe, E. L. Jones, D. B. Jones, P. Turyaguma, J. Oryema, D. H. Wright, and L. S. Young, 1995, Heterogeneous expression of Epstein-Barr virus latent proteins in endemic Burkitt's lymphoma: *Blood*, v. 86, p. 659 - 665.
- Niedobitek, G., L. S. Young, R. Lau, L. Brooks, D. Greenspan, J. S. Greenspan, and A. B. Rickinson, 1991, Epstein-Barr virus infection in oral hairy leukoplakia: virus replication in the absence of a detectable latent phase: *J Gen Virol*, v. 72 ( Pt 12), p. 3035-46.
- Nikitin, P. A., and M. A. Luftig, 2012, The DNA damage response in viral-induced cellular transformation: *British Journal of Cancer*, v. 106, p. 429-435.
- Nikitin, P. A., C. M. Yan, E. Forte, A. Bocedi, J. P. Tourigny, and R. E. White, 2010, An ATM/Chk2-mediated DNA damage-responsive signaling pathway suppresses



- Epstein-Barr virus transformation of primary human B cells: *Cell Host Microbe*, v. 8, p. 510-522.
- Nilsson, K., G. Klein, W. Henle, and G. Henle, 1971, The establishment of lymphoblastoid lines from adult and fetal human lymphoid tissue and its dependence on EBV: *Int J Cancer*, v. 8, p. 443-50.
- Ning, S., A. M. Hahn, L. E. Huye, and J. S. Pagano, 2003, Interferon regulatory factor 7 regulates expression of Epstein-Barr virus latent membrane protein 1: A regulatory circuit: *Journal of Virology*, v. 77, p. 9359-9368.
- Nolan, L. A., and A. J. Morgan, 1995, The Epstein-Barr virus open reading frame BDLF3 codes for a 100-150 kDa glycoprotein: *J Gen Virol*, v. 76 ( Pt 6), p. 1381-92.
- Oh, J.-K., and E. Weiderpass, 2014, Infection and Cancer: Global Distribution and Burden of Diseases: *Annals of Global Health*, v. 80, p. 384-392.
- Ok, C. Y., L. Li, and K. H. Young, 2015, EBV-driven B-cell lymphoproliferative disorders: from biology, classification and differential diagnosis to clinical management: *Exp Mol Med*, v. 47, p. e132.
- Okano, J., O. G. Opitz, H. Nakagawa, T. D. Jenkins, S. L. Friedman, and A. K. Rustgi, 2000, The Kruppel-like transcriptional factors Zf9 and GKLF coactivate the human keratin 4 promoter and physically interact: *FEBS Lett*, v. 473, p. 95-100.
- Old, L. J., E. A. Boyse, H. F. Oettgen, E. D. Harven, G. Geering, B. Williamson, and P. Clifford, 1966, Precipitating antibody in human serum to an antigen present in cultured Burkitt's lymphoma cells: *Proceedings of the National Academy of Sciences of the United States of America*, v. 56, p. 1699-1704.
- Ong, C. T., and V. G. Corces, 2011, Enhancer function: new insights into the regulation of tissue-specific gene expression: *Nat Rev Genet*, v. 12, p. 283-93.
- Pagano, J. S., 2009, EBV Diseases, *in* B. Damania, and J. M. Pipas, eds., *DNA Tumor Viruses*: New York, NY, Springer US, p. 217-240.
- Parker, B. D., A. Bankier, S. Satchwell, B. Barrell, and P. J. Farrell, 1990, Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B95-8 deletion region: *Virology*, v. 179, p. 339-46.
- Pathmanathan, R., U. Prasad, R. Sadler, K. Flynn, and N. Raab-Traub, 1995, Clonal Proliferations of Cells Infected with Epstein-Barr Virus in Preinvasive Lesions Related to Nasopharyngeal Carcinoma: *New England Journal of Medicine*, v. 333, p. 693-698.
- Patwardhan, R. P., J. B. Hiatt, D. M. Witten, M. J. Kim, R. P. Smith, D. May, C. Lee, J. M. Andrie, S. I. Lee, G. M. Cooper, N. Ahituv, L. A. Pennacchio, and J. Shendure, 2012, Massively parallel functional dissection of mammalian enhancers in vivo: *Nat Biotechnol*, v. 30, p. 265-70.
- Paulsen, S. J., M. M. Rosenkilde, J. Eugen-Olsen, and T. N. Kledal, 2005, Epstein-Barr virus-encoded BILF1 is a constitutively active G protein-coupled receptor: *J Virol*, v. 79, p. 536-46.
- Pellett, P. E., and B. Roizman, 2013, Herpesviridae, *In* Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, editors. (ed), *Fields virology*, v. vol 2: PA, Lippincott Williams & Wilkins, Philadelphia.
- Penner, C. G., and J. R. Davie, 1994, Transcription factor GATA-1-multiprotein complexes and chicken erythroid development: *FEBS Lett*, v. 342, p. 273-7.
- Petosa, C., P. Morand, F. Baudin, M. Moulin, J. B. Artero, and C. W. Müller, 2006, Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein: *Molecular Cell*, v. 21, p. 565-572.
- Pfeffer, S., A. Sewer, M. Lagos-Quintana, R. Sheridan, C. Sander, F. A. Grassler, L. F. van Dyk, C. K. Ho, S. Shuman, M. Chien, J. J. Russo, J. Ju, G. Randall, B. D. Lindenbach, C. M. Rice, V. Simon, D. D. Ho, M. Zavolan, and T. Tuschl, 2005, Identification of microRNAs of the herpesvirus family: *Nat Meth*, v. 2, p. 269-276.

- Pfeffer, S., M. Zavolan, F. A. Grasser, M. Chien, J. J. Russo, and J. Ju, 2004, Identification of virus-encoded microRNAs: *Science*, v. 304, p. 734-736.
- Piriou, E., A. S. Asito, P. O. Sumba, N. Fiore, J. M. Middeldorp, A. M. Moormann, R. Ploutz-Snyder, and R. Rochford, 2012, Early Age at Time of Primary Epstein-Barr Virus Infection Results in Poorly Controlled Viral Infection in Infants From Western Kenya: Clues to the Etiology of Endemic Burkitt Lymphoma: *The Journal of Infectious Diseases*, v. 205, p. 906-913.
- Portal, D., H. Zhou, B. Zhao, P. V. Kharchenko, E. Lowry, and L. Wong, 2013, Epstein-Barr virus nuclear antigen leader protein localizes to promoters and enhancers with cell transcription factors and EBNA2: *Proc Natl Acad Sci USA*, v. 110, p. 18537-18542.
- Pott, S., and J. D. Lieb, 2015, What are super-enhancers?: *Nat Genet*, v. 47, p. 8-12.
- Pratt, Z. L., M. Kuzembayeva, S. Sengupta, and B. Sugden, 2009, The microRNAs of Epstein-Barr Virus are expressed at dramatically differing levels among cell lines: *Virology*, v. 386, p. 387-397.
- Price, A. M., and M. A. Luftig, 2014, Dynamic Epstein-Barr virus gene expression on the path to B-cell transformation: *Adv Virus Res*, v. 88, p. 279-313.
- Quinn, L. L., L. R. Williams, C. White, C. Forrest, J. Zuo, and M. Rowe, 2016, The Missing Link in Epstein-Barr Virus Immune Evasion: the BDLF3 Gene Induces Ubiquitination and Downregulation of Major Histocompatibility Complex Class I (MHC-I) and MHC-II: *J Virol*, v. 90, p. 356-67.
- Quinn, L. L., J. Zuo, R. J. Abbott, C. Shannon-Lowe, R. J. Tierney, A. D. Hislop, and M. Rowe, 2014, Cooperation between Epstein-Barr virus immune evasion proteins spreads protection from CD8+ T cell recognition across all three phases of the lytic cycle: *PLoS Pathog*, v. 10, p. e1004322.
- Raab-Traub, N., 2009, Epstein-Barr Virus Transforming Proteins: Biologic Properties and Contribution to Oncogenesis, *in* B. Damania, and J. M. Pipas, eds., *DNA Tumor Viruses*: New York, NY, Springer US, p. 259-284.
- Raab-Traub, N., 2015, Nasopharyngeal Carcinoma: An Evolving Role for the Epstein-Barr Virus: *Curr Top Microbiol Immunol*, v. 390, p. 339-63.
- Raab-Traub, N., T. Dambaugh, and E. Kieff, 1980, DNA of Epstein-Barr virus VIII: B95-8, the previous prototype, is an unusual deletion derivative: *Cell*, v. 22, p. 257-67.
- Raab-Traub, N., and K. Flynn, 1986, The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation: *Cell*, v. 47, p. 883-889.
- Ragoczy, T., L. Heston, and G. Miller, 1998, The Epstein-Barr virus Rta protein disrupts latency in B lymphocytes: *J. Virol.*, v. 72, p. 7978-7984.
- Ragoczy, T., and G. Miller, 1999, Role of the Epstein-Barr Virus Rta Protein in Activation of Distinct Classes of Viral Lytic Cycle Genes: *Journal of Virology*, v. 73, p. 9858-9866.
- Rainey, J. J., W. O. Mwanda, P. Wairimu, A. M. Moormann, M. L. Wilson, and R. Rochford, 2007, Spatial distribution of Burkitt's lymphoma in Kenya and association with malaria risk: *Trop Med Int Health*, v. 12, p. 936 - 943.
- Ramasubramanyan, S., A. Kanhere, K. Osborn, K. Flower, R. G. Jenner, and A. J. Sinclair, 2012a, Genome-Wide Analyses of Zta Binding to the Epstein-Barr Virus Genome Reveals Interactions in both Early and Late Lytic Cycles and an Epigenetic Switch Leading to an Altered Binding Profile: *J Virol*, v. 86, p. 12494-502.
- Ramasubramanyan, S., K. Osborn, R. Al-Mohammad, I. B. Naranjo Perez-Fernandez, J. Zuo, N. Balan, A. Godfrey, H. Patel, G. Peters, M. Rowe, R. G. Jenner, and A. J. Sinclair, 2015, Epstein-Barr virus transcription factor Zta acts through distal regulatory elements to directly control cellular gene expression, *Nucleic Acids Res*, v. 43: England, The Author(s) 2015. Published by Oxford University Press on behalf of Nucleic Acids Research., p. 3563-77.

- Ramasubramanyan, S., K. Osborn, K. Flower, and A. J. Sinclair, 2012b, Dynamic Chromatin Environment of Key Lytic Cycle Regulatory Regions of the Epstein-Barr Virus Genome: *J Virol*, v. 86, p. 1809-19.
- Rani, C. S., N. Elango, S. S. Wang, K. Kobayashi, and R. Strong, 2009, Identification of an activator protein-1-like sequence as the glucocorticoid response element in the rat tyrosine hydroxylase gene: *Mol Pharmacol*, v. 75, p. 589-98.
- Reed, W., J. S. Carroll, and A. Agramonte, 1901, The etiology of yellow fever.: An additional note: *Journal of the American Medical Association*, v. XXXVI, p. 431-440.
- Reedman, B. M., and G. Klein, 1973, Cellular localization of an Epstein-Barr virus (EBV)-associated complement-fixing antigen in producer and non-producer lymphoblastoid cell lines: *Int J Cancer*, v. 11, p. 499-520.
- Rennekamp, A. J., and P. M. Lieberman, 2011, Initiation of Epstein-Barr Virus Lytic Replication Requires Transcription and the Formation of a Stable RNA-DNA Hybrid Molecule at OriLyt: *Journal of Virology*, v. 85, p. 2837-2850.
- Resnick, L., J. S. Herbst, D. V. Ablashi, and et al., 1988, REgression of oral hairy leukoplakia after orally administered acyclovir therapy: *JAMA*, v. 259, p. 384-388.
- Ressing, M. E., D. Horst, B. D. Griffin, J. Tellam, J. Zuo, R. Khanna, M. Rowe, and E. J. Wiertz, 2008, Epstein-Barr virus evasion of CD8(+) and CD4(+) T cell immunity via concerted actions of multiple gene products: *Semin Cancer Biol*, v. 18, p. 397-408.
- Ressing, M. E., S. E. Keating, D. van Leeuwen, D. Koppers-Lalic, I. Y. Pappworth, E. J. Wiertz, and M. Rowe, 2005a, Impaired transporter associated with antigen processing-dependent peptide transport during productive EBV infection: *J Immunol*, v. 174, p. 6829-38.
- Ressing, M. E., M. van Gent, A. M. Gram, M. J. Hooykaas, S. J. Piersma, and E. J. Wiertz, 2015, Immune Evasion by Epstein-Barr Virus: *Curr Top Microbiol Immunol*, v. 391, p. 355-81.
- Ressing, M. E., D. van Leeuwen, F. A. Verreck, R. Gomez, B. Heemskerk, M. Toebes, M. M. Mullen, T. S. Jardetzky, R. Longnecker, M. W. Schilham, T. H. Ottenhoff, J. Neefjes, T. N. Schumacher, L. M. Hutt-Fletcher, and E. J. Wiertz, 2003, Interference with T cell receptor-HLA-DR interactions by Epstein-Barr virus gp42 results in reduced T helper cell recognition: *Proc Natl Acad Sci U S A*, v. 100, p. 11583-8.
- Ressing, M. E., D. van Leeuwen, F. A. Verreck, S. Keating, R. Gomez, K. L. Franken, T. H. Ottenhoff, M. Spriggs, T. N. Schumacher, L. M. Hutt-Fletcher, M. Rowe, and E. J. Wiertz, 2005b, Epstein-Barr virus gp42 is posttranslationally modified to produce soluble gp42 that mediates HLA class II immune evasion: *J Virol*, v. 79, p. 841-52.
- Reusch, J. A., D. M. Nawandar, K. L. Wright, S. C. Kenney, and J. E. Mertz, 2015, Cellular differentiation regulator BLIMP1 induces Epstein-Barr virus lytic reactivation in epithelial and B cells by activating transcription from both the R and Z promoters: *J Virol*, v. 89, p. 1731-43.
- Rickinson, A. B., L. S. Young, and M. Rowe, 1987, Influence of the Epstein-Barr virus nuclear antigen EBNA 2 on the growth phenotype of virus-transformed B cells: *J Virol*, v. 61, p. 1310-7.
- Robinson, A. R., S. S. Kwek, and S. C. Kenney, 2012, The B-cell specific transcription factor, Oct-2, promotes Epstein-Barr virus latency by inhibiting the viral immediate-early protein, BZLF1: *PLoS Pathog*, v. 8, p. e1002516.
- Rochford, R., and A. M. Moormann, 2015, Burkitt's Lymphoma, *in* C. Münz, ed., *Epstein Barr Virus Volume 1: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 267-285.

- Rodriguez, A., E. J. Jung, Q. Yin, C. Cayrol, and E. K. Flemington, 2001, Role of c-myc regulation in Zta-mediated induction of the cyclin-dependent kinase inhibitors p21 and p27 and cell growth arrest: *Virology*, v. 284, p. 159-169.
- Roizman, B., L. E. Carmichael, F. Deinhardt, G. de-The, A. J. Nahmias, W. Plowright, F. Rapp, P. Sheldrick, M. Takahashi, and K. Wolf, 1981, Herpesviridae. Definition, provisional nomenclature, and taxonomy. The Herpesvirus Study Group, the International Committee on Taxonomy of Viruses: *Intervirology*, v. 16, p. 201-17.
- Roth, S., P. Franken, K. Monkhorst, J. Kong a San, and R. Fodde, 2012, Generation and characterization of an inducible transgenic model for studying mouse esophageal biology: *BMC Dev Biol*, v. 12, p. 18.
- Rott, R., and S. Siddell, 1998, One hundred years of animal virology: *J Gen Virol*, v. 79 ( Pt 11), p. 2871-4.
- Rous, P., 1911, A sarcoma of the fowl transmissible by an agent separable from the tumor cells: *The Journal of Experimental Medicine*, v. 13, p. 397-411.
- Rowe, M., B. Glaunsinger, D. van Leeuwen, J. Zuo, D. Sweetman, D. Ganem, J. Middeldorp, E. J. H. J. Wiertz, and M. E. Rensing, 2007, Host shutoff during productive Epstein–Barr virus infection is mediated by BGLF5 and may contribute to immune evasion: *Proceedings of the National Academy of Sciences*, v. 104, p. 3366-3371.
- Rymo, L., 1979, Identification of transcribed regions of Epstein-Barr virus DNA in Burkitt lymphoma-derived cells: *J Virol*, v. 32, p. 8-18.
- Røder, G., L. Geironsen, I. Bressendorff, and K. Paulsson, 2008, Viral Proteins Interfering with Antigen Presentation Target the Major Histocompatibility Complex Class I Peptide-Loading Complex: *Journal of Virology*, v. 82, p. 8246-8252.
- Sadler, R. H., and N. Raab-Traub, 1995, Structural analyses of the Epstein-Barr virus BamHI A transcripts: *J Virol*, v. 69, p. 1132-41.
- Salek-Ardakani, S., J. R. Arrand, and M. Mackett, 2002, Epstein–Barr Virus Encoded Interleukin-10 Inhibits HLA-Class I, ICAM-1, and B7 Expression on Human Monocytes: Implications for Immune Evasion by EBV: *Virology*, v. 304, p. 342-351.
- Samanta, M., D. Iwakiri, T. Kanda, T. Imaizumi, and K. Takada, 2006, EB virus-encoded RNAs are recognized by RIG-I and activate signaling to induce type I IFN: *EMBO J*, v. 25, p. 4207-4214.
- Sample, J. T., E. M. Marendy, D. J. Hughes, and C. E. Sample, 2009, The Epstein–Barr Virus Genome, *in* B. Damania, and J. M. Pipas, eds., *DNA Tumor Viruses*: New York, NY, Springer US, p. 241-258.
- Sander, S., D. P. Calado, L. Srinivasan, K. Kochert, B. Zhang, M. Rosolowski, S. J. Rodig, K. Holzmann, S. Stilgenbauer, R. Siebert, L. Bullinger, and K. Rajewsky, 2012, Synergy between PI3K signaling and MYC in Burkitt lymphomagenesis: *Cancer Cell*, v. 22, p. 167-79.
- Sandvej, K., L. Krenacs, S. J. Hamilton-Dutoit, J. L. Rindum, J. J. Pindborg, and G. Pallesen, 1992, Epstein-Barr virus latent and replicative gene expression in oral hairy leukoplakia: *Histopathology*, v. 20, p. 387-95.
- Sanyal, A., B. R. Lajoie, G. Jain, and J. Dekker, 2012, The long-range interaction landscape of gene promoters: *Nature*, v. 489, p. 109-13.
- Sarisky, R. T., Z. Gao, P. M. Lieberman, E. D. Fixman, G. S. Hayward, and S. D. Hayward, 1996, A replication function associated with the activation domain of the Epstein-Barr virus Zta transactivator: *J Virol*, v. 70, p. 8340-7.
- Sato, H., H. Takeshita, M. Furukawa, and M. Seiki, 1992, Epstein-Barr virus BZLF1 transactivator is a negative regulator of Jun: *Journal of Virology*, v. 66, p. 4732-4736.
- Schelcher, C., S. Al Mehairi, E. Verrall, Q. Hope, K. Flower, B. Bromley, D. N. Woolfson, M. J. West, and A. J. Sinclair, 2007, Atypical bZIP domain of viral



- transcription factor contributes to stability of dimer formation and transcriptional function: *Journal of Virology*, v. 81, p. 7149-7155.
- Schelcher, C., S. Valencia, H. J. Delecluse, M. Hicks, and A. J. Sinclair, 2005, Mutation of a Single Amino Acid Residue in the Basic Region of the Epstein-Barr Virus (EBV) Lytic Cycle Switch Protein Zta (BZLF1) Prevents Reactivation of EBV from Latency: *J Virol*, v. 79, p. 13822-8.
- Schepers, A., D. Pich, and W. Hammerschmidt, 1993a, A transcription factor with homology to the AP-1 family links RNA transcription and DNA replication in the lytic cycle of Epstein-Barr virus: *Embo j*, v. 12, p. 3921-9.
- Schepers, A., D. Pich, and W. Hammerschmidt, 1996, Activation of oriLyt, the lytic origin of DNA replication of Epstein-Barr virus, by BZLF1: *Virology*, v. 220, p. 367-76.
- Schepers, A., D. Pich, J. Mankertz, and W. Hammerschmidt, 1993b, cis-acting elements in the lytic origin of DNA replication of Epstein-Barr virus: *J Virol*, v. 67, p. 4237-45.
- Schirm, S., J. Jiricny, and W. Schaffner, 1987, The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity: *Genes Dev*, v. 1, p. 65-74.
- Schmitz, R., M. Ceribelli, S. Pittaluga, G. Wright, and L. M. Staudt, 2014, Oncogenic mechanisms in Burkitt lymphoma: *Cold Spring Harb Perspect Med*, v. 4.
- Segouffin-Cariou, C., G. Farjot, A. Sergeant, and H. Gruffat, 2000, Characterization of the Epstein-Barr virus BRRF1 gene, located between early genes BZLF1 and BRLF1: *J Gen Virol*, v. 81, p. 1791-9.
- Segre, J. A., C. Bauer, and E. Fuchs, 1999, Klf4 is a transcription factor required for establishing the barrier function of the skin: *Nat Genet*, v. 22, p. 356-360.
- Seibl, R., M. Motz, and H. Wolf, 1986, Strain-specific transcription and translation of the BamHI Z area of Epstein-Barr Virus: *Journal of Virology*, v. 60, p. 902-909.
- Serio, T. R., N. Cahill, M. E. Prout, and G. Miller, 1998, A functionally distinct TATA box required for late progression through the Epstein-Barr virus life cycle: *J Virol*, v. 72, p. 8338-43.
- Serio, T. R., J. L. Kolman, and G. Miller, 1997, Late gene expression from the Epstein-Barr virus BcLF1 and BFRF3 promoters does not require DNA replication in cis: *J Virol*, v. 71, p. 8726-34.
- Shaffer, A. L., K.-I. Lin, T. C. Kuo, X. Yu, E. M. Hurt, A. Rosenwald, J. M. Giltman, L. Yang, H. Zhao, K. Calame, and L. M. Staudt, 2002, Blimp-1 Orchestrates Plasma Cell Differentiation by Extinguishing the Mature B Cell Gene Expression Program: *Immunity*, v. 17, p. 51-62.
- Shannon-Lowe, C., E. Adland, A. I. Bell, H. J. Delecluse, A. B. Rickinson, and M. Rowe, 2009, Features distinguishing Epstein-Barr virus infections of epithelial cells and B cells: viral genome expression, genome maintenance, and genome amplification: *J Virol*, v. 83, p. 7749-60.
- Shannon-Lowe, C., and M. Rowe, 2011, Epstein-Barr virus infection of polarized epithelial cells via the basolateral surface by memory B cell-mediated transfer infection: *PLoS Pathog*, v. 7, p. e1001338.
- Shannon-Lowe, C. D., B. Neuhierl, G. Baldwin, A. B. Rickinson, and H. J. Delecluse, 2006, Resting B cells as a transfer vehicle for Epstein-Barr virus infection of epithelial cells: *Proc Natl Acad Sci U S A*, v. 103, p. 7065-70.
- Sharon, E., Y. Kalma, A. Sharp, T. Raveh-Sadka, M. Levo, D. Zeevi, L. Keren, Z. Yakhini, A. Weinberger, and E. Segal, 2012, Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters: *Nat Biotechnol*, v. 30, p. 521-30.
- Shen, Y., F. Yue, D. F. McCleary, Z. Ye, L. Edsall, S. Kuan, U. Wagner, J. Dixon, L. Lee, V. V. Lobanenko, and B. Ren, 2012, A map of the cis-regulatory sequences in the mouse genome: *Nature*, v. 488, p. 116-20.

- Shie, J.-L., Z. Y. Chen, M. Fu, R. G. Pestell, and C.-C. Tseng, 2000, Gut-enriched Krüppel-like factor represses cyclin D1 promoter activity through Sp1 motif: *Nucleic Acids Research*, v. 28, p. 2969-2976.
- Shlyueva, D., G. Stampfel, and A. Stark, 2014, Transcriptional enhancers: from properties to genome-wide predictions: *Nat Rev Genet*, v. 15, p. 272-286.
- Sinclair, A. J., 2003, bZIP proteins of human gammaherpesviruses: *Journal of General Virology*, v. 84, p. 1941-1949.
- Sinclair, A. J., 2006, Unexpected structure of Epstein-Barr virus lytic cycle activator Zta: *Trends in Microbiology*, v. 14, p. 289-291.
- Sinclair, A. J., M. Brimmell, F. Shanahan, and P. J. Farrell, 1991, Pathways of activation of the Epstein-Barr virus productive cycle: *J Virol*, v. 65, p. 2237-44.
- Sinclair, A. J., and P. J. Farrell, 1995, Host cell requirements for efficient infection of quiescent primary B lymphocytes by Epstein-Barr virus: *Journal of Virology*, v. 69, p. 5461-5468.
- Sinclair, A. J., I. Palmero, G. Peters, and P. J. Farrell, 1994a, EBNA-2 and EBNA-LP cooperate to cause G0 and G1 transition during immortalization of resting human B lymphocytes by Epstein-Barr virus: *EMBO J.*, v. 13, p. 3321-3328.
- Sinclair, A. J., I. Palmero, G. Peters, and P. J. Farrell, 1994b, EBNA-2 and EBNA-LP cooperate to cause G0 to G1 transition during immortalization of resting human B lymphocytes by Epstein-Barr virus: *Embo j*, v. 13, p. 3321-8.
- Sixbey, J. W., J. G. Nedrud, N. Raab-Traub, R. A. Hanes, and J. S. Pagano, 1984, Epstein-Barr virus replication in oropharyngeal epithelial cells: *N Engl J Med*, v. 310, p. 1225-30.
- Skalsky, R. L., and B. R. Cullen, 2015, EBV Noncoding RNAs, *in* C. Münz, ed., *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*: Cham, Springer International Publishing, p. 181-217.
- Smale, S. T., and J. T. Kadonaga, 2003, The RNA polymerase II core promoter: *Annu Rev Biochem*, v. 72, p. 449-79.
- Smets, F., D. Latinne, H. Bazin, R. Reding, J. B. Otte, J. P. Buts, and E. M. Sokal, 2002, Ratio between Epstein-Barr viral load and anti-Epstein-Barr virus specific T-cell response as a predictive marker of posttransplant lymphoproliferative disease: *Transplantation*, v. 73, p. 1603-10.
- Smith, P. R., 2000, Structure and coding content of CST (BART) family RNAs of Epstein-Barr virus: *J. Virol.*, v. 74, p. 3082-3092.
- Smith, P. R., O. de Jesus, D. Turner, M. Hollyoake, C. E. Karstegl, B. E. Griffin, L. Karran, Y. Wang, S. D. Hayward, and P. J. Farrell, 2000, Structure and coding content of CST (BART) family RNAs of Epstein-Barr virus: *J Virol*, v. 74, p. 3082-92.
- Soufi, A., M. F. Garcia, A. Jaroszewicz, N. Osman, M. Pellegrini, and K. S. Zaret, 2015, Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming: *Cell*, v. 161, p. 555-568.
- Spitz, F., and E. E. Furlong, 2012, Transcription factors: from enhancer binding to developmental control: *Nat Rev Genet*, v. 13, p. 613-26.
- Stolzenberg, M. C., and T. Ooka, 1990, Purification and properties of Epstein-Barr virus DNase expressed in *Escherichia coli*: *J Virol*, v. 64, p. 96-104.
- Stone, M. J., 2005, Thomas Hodgkin: medical immortal and uncompromising idealist: *Proceedings (Baylor University. Medical Center)*, v. 18, p. 368-375.
- Strong, M. J., T. Laskow, H. Nakhoul, E. Blanchard, Y. Liu, X. Wang, M. Baddoo, Z. Lin, Q. Yin, and E. K. Flemington, 2015, Latent Expression of the Epstein-Barr Virus (EBV)-Encoded Major Histocompatibility Complex Class I TAP Inhibitor, BNLF2a, in EBV-Positive Gastric Carcinomas: *J Virol*, v. 89, p. 10110-4.
- Sur, I., and J. Taipale, 2016, The role of enhancers in cancer: *Nat Rev Cancer*, v. 16, p. 483-493.
- Takada, K., 2000, Epstein-Barr virus and gastric carcinoma: *Molecular Pathology*, v. 53, p. 255-261.

- Takada, K., and Y. Ono, 1989, Synchronous and sequential activation of latently infected Epstein-Barr virus genomes: *Journal of Virology*, v. 63, p. 445-449.
- Takada, K., N. Shimizu, S. Sakuma, and Y. Ono, 1986, trans activation of the latent Epstein-Barr virus (EBV) genome after transfection of the EBV DNA fragment: *J Virol*, v. 57, p. 1016-22.
- Tanner, J. E., and C. Alfieri, 2001, The Epstein-Barr virus and post-transplant lymphoproliferative disease: interplay of immunosuppression, EBV, and the immune system in disease pathogenesis: *Transpl Infect Dis*, v. 3, p. 60-9.
- Taylor, N., E. Flemington, J. L. Kolman, R. P. Baumann, S. H. Speck, and G. Miller, 1991, ZEBRA and a Fos-GCN4 chimeric protein differ in their DNA-binding specificities for sites in the Epstein-Barr virus BZLF1 promoter: *Journal of Virology*, v. 65, p. 4033-4041.
- Temple, R. M., J. Zhu, L. Budgeon, N. D. Christensen, C. Meyers, and C. E. Sample, 2014, Efficient replication of Epstein-Barr virus in stratified epithelium in vitro: *Proceedings of the National Academy of Sciences*, v. 111, p. 16544-16549.
- Thierry, F., G. Spyrou, M. Yaniv, and P. Howley, 1992, Two AP1 sites binding JunB are essential for human papillomavirus type 18 transcription in keratinocytes: *Journal of Virology*, v. 66, p. 3740-3748.
- Thomas, C., A. Dankesreiter, H. Wolf, and F. Schwarzmann, 2003, The BZLF1 promoter of Epstein-Barr virus is controlled by E box-/HI-motif-binding factors during virus latency: *Journal of General Virology*, v. 84, p. 959-964.
- Thorley-Lawson, D. A., 2015, EBV Persistence--Introducing the Virus: *Curr Top Microbiol Immunol*, v. 390, p. 151-209.
- Thorley-Lawson, D. A., J. B. Hawkins, S. I. Tracy, and M. Shapiro, 2013, The pathogenesis of Epstein-Barr virus persistent infection: *Curr Opin Virol*, v. 3, p. 227-32.
- Tidy, H., 1950, Glandular Fever: Infectious Mononucleosis: *Postgraduate Medical Journal*, v. 26, p. 9-15.
- Tierney, R. J., C. D. Shannon-Lowe, L. Fitzsimmons, A. I. Bell, and M. Rowe, 2015, Unexpected patterns of Epstein-Barr virus transcription revealed by a High throughput PCR array for absolute quantification of viral mRNA: *Virology*, v. 474, p. 117-130.
- Tsai, M. H., A. Raykova, O. Klinke, K. Bernhardt, K. Gartner, C. S. Leung, K. Geletneky, S. Sertel, C. Munz, R. Feederle, and H. J. Delecluse, 2013, Spontaneous lytic replication and epitheliotropism define an Epstein-Barr virus strain found in carcinomas: *Cell Rep*, v. 5, p. 458-70.
- Tsai, S. C., S. J. Lin, P. W. Chen, W. Y. Luo, T. H. Yeh, H. W. Wang, C. J. Chen, and C. H. Tsai, 2009, EBV Zta protein induces the expression of interleukin-13, promoting the proliferation of EBV-infected B cells and lymphoblastoid cell lines: *Blood*, v. 114, p. 109-18.
- Tse, E., and Y.-L. Kwong, 2013, How I treat NK/T-cell lymphomas: *Blood*, v. 121, p. 4997.
- Tsuchiya, T., M. Okada, M. Ueda, and Y. Yasukochi, 1997, Activation of the erythropoietin promoter by a point mutation from GATA to TATA in the -30 region: *J Biochem*, v. 121, p. 193-6.
- Tugizov, S., R. Herrera, P. Veluppillai, J. Greenspan, D. Greenspan, and J. M. Palefsky, 2007, Epstein-Barr Virus (EBV)-Infected Monocytes Facilitate Dissemination of EBV within the Oral Mucosal Epithelium: *Journal of Virology*, v. 81, p. 5484-5496.
- Urier, G., M. Buisson, P. Chambard, and A. Sergeant, 1989, The Epstein-Barr virus early protein EB1 activates transcription from different responsive elements including AP-1 binding sites: *Embo j*, v. 8, p. 1447-53.
- van Gent, M., B. D. Griffin, E. G. Berkhoff, D. van Leeuwen, I. G. Boer, M. Buisson, F. C. Hartgers, W. P. Burmeister, E. J. Wiertz, and M. E. Rensing, 2011, EBV lytic-

- phase protein BGLF5 contributes to TLR9 downregulation during productive infection: *J Immunol*, v. 186, p. 1694-702.
- Visel, A., S. Minovitsky, I. Dubchak, and L. A. Pennacchio, 2007, VISTA Enhancer Browser--a database of tissue-specific human enhancers: *Nucleic Acids Res*, v. 35, p. D88-92.
- Vliet, J. v., J. Turner, and M. Crossley, 2000, Human Krüppel-like Factor 8: a CACCC-box binding protein that associates with CtBP and represses transcription: *Nucleic Acids Research*, v. 28, p. 1955-1962.
- Walling, D. M., C. M. Flaitz, and C. M. Nichols, 2003, Epstein-Barr Virus Replication in Oral Hairy Leukoplakia: Response, Persistence, and Resistance to Treatment with Valacyclovir: *Journal of Infectious Diseases*, v. 188, p. 883-890.
- Wang, P., L. Day, and P. M. Lieberman, 2006, Multivalent Sequence Recognition by Epstein-Barr Virus Zta Requires Cysteine 171 and an Extension of the Canonical B-ZIP Domain: *Journal of Virology*, v. 80, p. 10942-10949.
- Webster-Cyriaque, J., J. Middeldorp, and N. Raab-Traub, 2000, Hairy Leukoplakia: an Unusual Combination of Transforming and Permissive Epstein-Barr Virus Infections: *Journal of Virology*, v. 74, p. 7610-7618.
- Wen, W., D. Iwakiri, K. Yamamoto, S. Maruo, T. Kanda, and K. Takada, 2007, Epstein-Barr Virus BZLF1 Gene, a Switch from Latency to Lytic Infection, Is Expressed as an Immediate-Early Gene after Primary Infection of B Lymphocytes: *Journal of Virology*, v. 81, p. 1037-1042.
- Whyte, W. A., D. A. Orlando, D. Hnisz, B. J. Abraham, C. Y. Lin, M. H. Kagey, P. B. Rahl, T. I. Lee, and R. A. Young, 2013, Master transcription factors and mediator establish super-enhancers at key cell identity genes: *Cell*, v. 153, p. 307-19.
- Wiedmer, A., P. Wang, J. Zhou, A. J. Rennekamp, V. Tiranti, M. Zeviani, and P. M. Lieberman, 2008, Epstein-Barr virus immediate-early protein Zta co-opts mitochondrial single-stranded DNA binding protein to promote viral and inhibit mitochondrial DNA replication: *J Virol*, v. 82, p. 4647-55.
- Wille, C. K., D. M. Nawandar, A. R. Panfil, M. M. Ko, S. R. Hagemeier, and S. C. Kenney, 2013, Viral genome methylation differentially affects the ability of BZLF1 versus BRLF1 to activate Epstein-Barr virus lytic gene expression and viral replication: *J Virol*, v. 87, p. 935-50.
- Wilson, J. B., J. L. Bell, and A. J. Levine, 1996, Expression of Epstein-Barr virus nuclear antigen-1 induces B cell neoplasia in transgenic mice: *EMBO J.*, v. 15, p. 3117-3126.
- Wilson, J. B., W. Weinberg, R. Johnson, S. Yuspa, and A. J. Levine, 1990, Expression of the BNLF-1 oncogene of Epstein-Barr virus in the skin of transgenic mice induces hyperplasia and aberrant expression of keratin 6: *Cell*, v. 61, p. 1315-27.
- Woellmer, A., J. M. Arteaga-Salas, and W. Hammerschmidt, 2012, BZLF1 Governs CpG-Methylated Chromatin of Epstein-Barr Virus Reversing Epigenetic Repression: *PLoS Pathog*, v. 8, p. e1002902.
- Woellmer, A., and W. Hammerschmidt, 2013, Epstein-Barr virus and host cell methylation: Regulation of latency, replication and virus reactivation: *Current opinion in virology*, v. 3, p. 260-265.
- Wolf, H., J. Werner, and H. zur Hausen, 1975, EBV DNA in nonlymphoid cells of nasopharyngeal carcinomas and in a malignant lymphoma obtained after inoculation of EBV into cottontop marmosets: *Cold Spring Harb Symp Quant Biol*, v. 39 Pt 2, p. 791-6.
- Wu, F. Y., H. Chen, S. E. Wang, C. M. J. ApRhys, G. Liao, M. Fujimuro, C. J. Farrell, J. Huang, S. D. Hayward, and G. S. Hayward, 2003, CCAAT/enhancer binding protein  $\alpha$  interacts with ZTA and mediates ZTA-induced p21CIP-1 accumulation and G1 cell cycle arrest during the Epstein-Barr virus lytic cycle: *Journal of Virology*, v. 77, p. 1481-1500.



- Wu, F. Y., S. E. Wang, H. Chen, L. Wang, S. D. Hayward, and G. S. Hayward, 2004, CCAAT/enhancer binding protein alpha binds to the Epstein-Barr virus (EBV) ZTA protein through oligomeric interactions and contributes to cooperative transcriptional activation of the ZTA promoter through direct binding to the ZII and ZIIIB motifs during induction of the EBV lytic cycle: *J Virol*, v. 78, p. 4847-65.
- Wycisk, A. I., J. Lin, S. Loch, K. Hobohm, J. Funke, R. Wieneke, J. Koch, W. R. Skach, P. U. Mayerhofer, and R. Tampe, 2011, Epstein-Barr viral BNLF2a protein hijacks the tail-anchored protein insertion machinery to block antigen processing by the transport complex TAP: *J Biol Chem*, v. 286, p. 41402-12.
- Xiao, H., J. D. Friesen, and J. T. Lis, 1995, Recruiting TATA-binding protein to a promoter: transcriptional activation without an upstream activator: *Mol Cell Biol*, v. 15, p. 5757-61.
- Yajima, M., T. Kanda, and K. Takada, 2005, Critical role of Epstein-Barr Virus (EBV)-encoded RNA in efficient EBV-induced B-lymphocyte growth transformation: *J Virol*, v. 79, p. 4298-4307.
- Yang, H., K. Yu, R. Zhang, J. Li, X. Wei, Y. Zhang, C. Zhang, F. Xiao, D. Zhao, X. Lin, H. Wu, and X. Yang, 2016, The HLA-DRB1 allele polymorphisms and nasopharyngeal carcinoma: *Tumor Biology*, v. 37, p. 7119-7128.
- Yang, H. J., T. J. Huang, C. F. Yang, L. X. Peng, R. Y. Liu, G. D. Yang, Q. Q. Chu, J. L. Huang, N. Liu, H. B. Huang, Z. Y. Zhu, C. N. Qian, and B. J. Huang, 2013, Comprehensive profiling of Epstein-Barr virus-encoded miRNA species associated with specific latency types in tumor cells: *Virol J*, v. 10, p. 314.
- Yang, W. T., and P. S. Zheng, 2012, Kruppel-like factor 4 functions as a tumor suppressor in cervical carcinoma: *Cancer*, v. 118, p. 3691-702.
- Yang, X., J. S. Sham, M. H. Ng, S. W. Tsao, D. Zhang, S. W. Lowe, and L. Cao, 2000, LMP1 of Epstein-Barr virus induces proliferation of primary mouse embryonic fibroblasts and cooperatively transforms the cells with a p16-insensitive CDK4 oncogene: *J Virol*, v. 74, p. 883-91.
- Yang, Y. C., T. H. Feng, T. Y. Chen, H. H. Huang, C. C. Hung, S. T. Liu, and L. K. Chang, 2015, RanBPM regulates Zta-mediated transcriptional activity in Epstein-Barr virus: *J Gen Virol*, v. 96, p. 2336-48.
- Yates, J., N. Warren, D. Reisman, and B. Sugden, 1984, A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells: *Proc Natl Acad Sci USA*, v. 81, p. 3806-3810.
- Yates, J. L., S. M. Camiolo, and J. M. Bashaw, 2000, The minimal replicator of Epstein-Barr virus oriP: *J Virol*, v. 74, p. 4512-22.
- Yori, J. L., E. Johnson, G. Zhou, M. K. Jain, and R. A. Keri, 2010, Kruppel-like factor 4 inhibits epithelial-to-mesenchymal transition through regulation of E-cadherin gene expression: *J Biol Chem*, v. 285, p. 16854-63.
- Young, L. S., and C. W. Dawson, 2014, Epstein-Barr virus and nasopharyngeal carcinoma: *Chinese Journal of Cancer*, v. 33, p. 581-590.
- Young, L. S., R. Lau, M. Rowe, G. Niedobitek, G. Packham, and F. Shanahan, 1991a, Differentiation-associated expression of the Epstein-Barr virus BZLF1 transactivator protein in oral hairy leukoplakia: *J Virol*, v. 65, p. 2868-2874.
- Young, L. S., R. Lau, M. Rowe, G. Niedobitek, G. Packham, F. Shanahan, D. T. Rowe, D. Greenspan, J. S. Greenspan, and A. B. Rickinson, 1991b, Differentiation-associated expression of the Epstein-Barr virus BZLF1 transactivator protein in oral hairy leukoplakia: *Journal of Virology*, v. 65, p. 2868-2874.
- Young, L. S., R. Lau, M. Rowe, G. Niedobitek, G. Packham, F. Shanahan, D. T. Rowe, D. Greenspan, J. S. Greenspan, and A. B. Rickinson, 1991c, Differentiation-associated expression of the Epstein-Barr virus BZLF1 transactivator protein in oral hairy leukoplakia: *J Virol*, v. 65, p. 2868-74.
- Young, L. S., and P. G. Murray, 2003, Epstein-Barr virus and oncogenesis: from latent genes to tumours: *Oncogene*, v. 22, p. 5108-5121.

- Young, L. S., and A. B. Rickinson, 2004, Epstein-Barr virus: 40 years on: *Nat Rev Cancer*, v. 4, p. 757-768.
- Yu, M. C., J. H. C. Ho, R. K. Ross, and B. E. Henderson, 1981, Nasopharyngeal carcinoma in Chinese—Salted fish or inhaled smoke?: *Preventive Medicine*, v. 10, p. 15-24.
- Yuan, J., E. Cahir-McFarland, B. Zhao, and E. Kieff, 2006, Virus and cell RNAs expressed during Epstein-Barr virus replication: *J Virol*, v. 80, p. 2548-65.
- Zaret, K. S., and J. S. Carroll, 2011, Pioneer transcription factors: establishing competence for gene expression: *Genes Dev*, v. 25, p. 2227-41.
- Zech, L., U. Haglund, K. Nilsson, and G. Klein, 1976, Characteristic chromosomal abnormalities in biopsies and lymphoid-cell lines from patients with Burkitt and non-Burkitt lymphomas: *Int J Cancer*, v. 17, p. 47-56.
- Zerby, D., C. J. Chen, E. Poon, D. Lee, R. Shiekhatar, and P. M. Lieberman, 1999, The amino-terminal C/H1 domain of CREB binding protein mediates zta transcriptional activation of latent Epstein-Barr virus: *Mol Cell Biol*, v. 19, p. 1617-26.
- Zhang, L., Q.-Y. Chen, H. Liu, L.-Q. Tang, and H.-Q. Mai, 2013, Emerging treatment options for nasopharyngeal carcinoma: *Drug Design, Development and Therapy*, v. 7, p. 37-52.
- Zhang, Q., D. Gutsch, and S. Kenney, 1994, Functional and physical interaction between p53 and BZLF1: implications for Epstein-Barr virus latency: *Mol Cell Biol*, v. 14, p. 1929-38.
- Zhang, Q., Y. Hong, D. Dorsky, E. Holley-Guthrie, S. Zalani, N. A. Elshiekh, A. Kiehl, T. Le, and S. Kenney, 1996, Functional and physical interactions between the Epstein-Barr virus (EBV) proteins BZLF1 and BMRF1: Effects on EBV transcription and lytic replication: *J Virol*, v. 70, p. 5131-42.
- Zhang, X. L., A. Langford, J. Becker, J. P. Rabanus, H. D. Pohle, P. Reichart, and H. Gelderblom, 1988, Ultrastructural and immunohistochemical findings in oral hairy leukoplakia: *Virchows Arch A Pathol Anat Histopathol*, v. 412, p. 533-42.
- Zhou, H., S. C. Schmidt, S. Jiang, B. Willox, K. Bernhardt, J. Liang, E. C. Johannsen, P. Kharchenko, B. E. Gewurz, E. Kieff, and B. Zhao, 2015, Epstein-Barr virus oncoprotein super-enhancers control B cell growth: *Cell Host Microbe*, v. 17, p. 205-16.
- Zhou, J., C. M. Chau, Z. Deng, R. Shiekhatar, M. P. Spindler, A. Schepers, and P. M. Lieberman, 2005, Cell cycle regulation of chromatin at an origin of DNA replication: *Embo j*, v. 24, p. 1406-17.
- Zhu, H., G. Wang, and J. Qian, 2016, Transcription factors as readers and effectors of DNA methylation: *Nat Rev Genet*, v. 17, p. 551-565.
- Zuniga, E. I., M. Macal, G. M. Lewis, and J. A. Harker, 2015, Innate and Adaptive Immune Regulation During Chronic Viral Infections: *Annual review of virology*, v. 2, p. 573-597.
- Zuo, J., A. Currin, B. D. Griffin, C. Shannon-Lowe, W. A. Thomas, M. E. Rensing, E. J. H. J. Wiertz, and M. Rowe, 2009, The Epstein-Barr Virus G-Protein-Coupled Receptor Contributes to Immune Evasion by Targeting MHC Class I Molecules for Degradation: *PLoS Pathog*, v. 5, p. e1000255.
- Zuo, J., L. L. Quinn, J. Tamblyn, W. A. Thomas, R. Feederle, H.-J. Delecluse, A. D. Hislop, and M. Rowe, 2011a, The Epstein-Barr Virus-Encoded BILF1 Protein Modulates Immune Recognition of Endogenously Processed Antigen by Targeting Major Histocompatibility Complex Class I Molecules Trafficking on both the Exocytic and Endocytic Pathways: *Journal of Virology*, v. 85, p. 1604-1614.
- Zuo, J., W. Thomas, D. van Leeuwen, J. M. Middeldorp, E. J. Wiertz, M. E. Rensing, and M. Rowe, 2008, The DNase of gammaherpesviruses impairs recognition by virus-specific CD8<sup>+</sup> T cells through an additional host shutoff function: *J Virol*, v. 82, p. 2385-93.

- Zuo, J., W. A. Thomas, T. A. Haigh, L. Fitzsimmons, H. M. Long, A. D. Hislop, G. S. Taylor, and M. Rowe, 2011b, Epstein-Barr virus evades CD4+ T cell responses in lytic cycle through BZLF1-mediated downregulation of CD74 and the cooperation of vBcl-2: PLoS Pathog, v. 7, p. e1002455.
- zur Hausen, H., 1972, Epstein-Barr virus in human tumor cells: Int Rev Exp Pathol, v. 11, p. 233-58.
- zur Hausen, H., H. Schulte-Holthausen, G. Klein, W. Henle, G. Henle, P. Clifford, and L. Santesson, 1970, EBV DNA in biopsies of Burkitt tumours and anaplastic carcinomas of the nasopharynx: Nature, v. 228, p. 1056-8.

## **Appendix A. Publications**



# Epstein–Barr virus transcription factor Zta acts through distal regulatory elements to directly control cellular gene expression

Sharada Ramasubramanian<sup>1,†</sup>, Kay Osborn<sup>1,†</sup>, Rajaei Al-Mohammad<sup>1,†</sup>, Ijiel B. Naranjo Perez-Fernandez<sup>1,†</sup>, Jianmin Zuo<sup>2</sup>, Nicolae Balan<sup>1</sup>, Anja Godfrey<sup>1</sup>, Harshil Patel<sup>3</sup>, Gordon Peters<sup>3</sup>, Martin Rowe<sup>2</sup>, Richard G. Jenner<sup>4</sup> and Alison J. Sinclair<sup>1,\*</sup>

<sup>1</sup>School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK, <sup>2</sup>School of Cancer Sciences, The University of Birmingham, Birmingham B15 2TT, UK, <sup>3</sup>Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3LY, UK and <sup>4</sup>UCL Cancer Institute and MRC Centre for Medical Molecular Virology, Paul O'Gorman Building, University College London, London W1CE 6BT, UK

Received November 23, 2014; Revised February 23, 2015; Accepted March 01, 2015

## ABSTRACT

**Lytic replication of the human gamma herpes virus Epstein-Barr virus (EBV) is an essential prerequisite for the spread of the virus. Differential regulation of a limited number of cellular genes has been reported in B-cells during the viral lytic replication cycle. We asked whether a viral bZIP transcription factor, Zta (BZLF1, ZEBRA, EB1), drives some of these changes. Using genome-wide chromatin immunoprecipitation coupled to next-generation DNA sequencing (ChIP-seq) we established a map of Zta interactions across the human genome. Using sensitive transcriptome analyses we identified 2263 cellular genes whose expression is significantly changed during the EBV lytic replication cycle. Zta binds 278 of the regulated genes and the distribution of binding sites shows that Zta binds mostly to sites that are distal to transcription start sites. This differs from the prevailing view that Zta activates viral genes by binding exclusively at promoter elements. We show that a synthetic Zta binding element confers Zta regulation at a distance and that distal Zta binding sites from cellular genes can confer Zta-mediated regulation on a heterologous promoter. This leads us to propose that Zta directly reprograms the expression of cellular genes through distal elements.**

## INTRODUCTION

Epstein–Barr virus represents a serious health threat both in the developing world and in western economies and is a sig-

nificant risk factor for Hodgkin's lymphoma (HL), Burkitt's lymphoma (BL) and nasopharyngeal carcinoma (NPC) (1–4). Primary infection with EBV is also responsible for the development of infectious mononucleosis (5). Infection of resting primary B-lymphocytes by EBV generates a population of cells that are effectively immortal. This represents the first step in the establishment of life-long viral latency *in vivo* and generates precursors that can develop into lymphoid malignancies (6,7).

EBV genes fall primarily into two groups, depending on their patterns of expression, and are classified as either latent or lytic (6,7). During latency, the majority of the viral genome is silenced and only a restricted set of genes is expressed (6,7). Once EBV lytic cycle replication is initiated, the lytic cycle genes are transcriptionally activated (8,9). Some lytic cycle genes are also expressed following infection of B-lymphocytes by EBV, but this does not result in a full lytic replication cycle (10–13) and has been termed a pre-latency phase (14).

The switch from viral latency to the lytic replication cycle is critical for virus spread. A key executioner of this process is the transcription and replication factor Zta (BZLF1, ZEBRA, EB1, Z) (9,15). This AP-1-like viral protein is a member of the bZIP family and interacts with a seven base-pair DNA sequence element termed Zta-response element (ZRE), for which at least 32 sequence variants are known (16,17). A sub-set of ZREs contain a CpG motif and are only recognized as ZREs when the DNA is methylated (11,18,19). DNA binding analyses across the EBV genome by ourselves and others revealed that Zta interacts extensively with viral promoters and with the viral origins of lytic replication (16,20). The current model of Zta function is

\*To whom correspondence should be addressed. Tel: +44 1273 678 194; Email: a.j.sinclair@sussex.ac.uk

<sup>†</sup>The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors and that the third and fourth author contributed equally.

that it acts at promoter-proximal regions to recruit RNA polymerase II and associated factors (15).

As well as regulating viral genes, Zta has the potential to alter the patterns of gene expression in the host cell. Zta is known to regulate the expression of a few cellular genes such as *FOS* (21), *E2F1* [23], *EGRI* (22,23), *IL8* (24), *IL10* (25) and *IL13* (26). Furthermore, attempts have been made to obtain a global map of cellular genes regulated during EBV replication using an EBV-positive BL cell line induced to initiate viral replication (27,28). However, as only a minority of the cells in a given population responds to the activating stimulus, this approach detected only the most highly activated genes.

Here, we undertook an unbiased genome-wide survey to map cellular genes that are bound by Zta. We then generated a cell system that allows enrichment of cells undergoing lytic cycle to perform sensitive transcriptome analyses of changes in expression of cellular genes during the lytic cycle. Integrating these data-sets led us to propose that Zta is able to regulate the expression of cellular genes through interactions with distal elements as well as through promoters, and we present further evidence to support this hypothesis.

## MATERIALS AND METHODS

### Cell lines

The Akata group I EBV-positive BL cell line (29), an EBV-immortalized lymphoblastoid cell line, LCL#3 (30) and an EBV-negative BL cell line, DG75 (31), were maintained in RPMI medium supplemented with 10% (v/v) fetal bovine serum, 100 units/ml penicillin, 100 µg/ml streptomycin and 2 mM L-glutamine (Life Technologies) at 37°C with 5% CO<sub>2</sub>. For EBV lytic induction, Akata cells were seeded in log phase growth at  $5 \times 10^5$  cells/ml. After 24 h, the cells were concentrated to  $2 \times 10^6$  cells/ml and treated with 0.125% rabbit anti-human IgG (DAKO) or Dulbecco's phosphate buffered saline (DPBS) for the indicated times, in the presence or absence of acyclovir (Sigma) to capture both early and late lytic stages of viral replication (20). HONE1-EBV cells were maintained in the same medium and the presence of EBV within the cells was maintained by selection with 600 µg/ml G418. For EBV lytic replication cycle induction, HONE1-EBV cells were grown to 70% confluence and induced with 10 µM suberoylanilide hydroxamic acid (SAHA) for 48 h.

The expression vector pRTS-CD2-BZLF1, which drives expression of Zta, non-functional NGFR and GFP from a bi-directional doxycycline-regulated promoter and CD2 from a constitutive promoter or a control vector in which the BZLF1 sequences are cloned in the reverse orientation (32), were introduced into the Akata cell line by electroporation. Cells harboring the plasmids were physically enriched on the basis of CD2 expression, generating Akata-Zta and Akata control cells as described (32). Expression of the doxycycline-regulated promoter was induced by addition of 500 ng/ml doxycycline for 4–24 h and NGFR-expressing cells were isolated with anti-NGFR antibodies coupled to paramagnetic beads as described (32). Viable cell counts were assessed by trypan blue (Sigma) exclusion.

### FACS and protein analyses

GFP-positive cells were detected using multi-parameter fluorescent activated cell analysis (FACS) (Facs Canto-Beckton Dickinson). Intracellular staining of Zta was undertaken with the mouse monoclonal antibody BZ1 (33), with anti-mouse IgG coupled to Donkey F(ab')<sub>2</sub> as a secondary reagent (anti-mouse IgG -Alexa Fluor® 647) and the FIX & PERM® Cell Permeabilization Kit (Life Technologies). Intracellular staining of VCA was undertaken with the mouse monoclonal antibody VCA-gp125 (clone L2 MAB8184, Millipore) and the same secondary reagent and conditions. Dual parameters of GFP and Alexa-fluor 647-coupled staining were collected and the double positive cells identified using BD FACSDiva™ Software (Beckton Dickinson).

Cells were harvested in sample buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue and 0.125 M Tris-HCl, pH 6.8.) and total proteins were fractionated by SDS-PAGE in a 12% gel (Novex). After transfer to nitrocellulose membrane, Zta and actin were detected by immunoblotting with the BZ1 mouse monoclonal antibody against Zta (33) and a rabbit antibody against β-actin (Sigma). Species-specific IR-labeled anti-mouse and anti-rabbit antibodies (Licor) were used in the second layer and the resulting signal was detected on an Odyssey Fc Imager and quantified using Odyssey Image Studio (Licor).

### Luciferase reporter assays

A 239 bp DNA element comprising 20 non-CpG ZREs (17) with a spacing element of 4–6 nucleotides between each was synthesized as complementary oligonucleotides (Gene Strings—Life Technologies) and inserted between the SalI and BamHI sites at the enhancer cloning site of the pGL3-enhancer reporter vector (Promega). Two minimal promoters, one from the cell CIITA gene (−214/+54) and the minimal promoter from the pGL4.23 reporter vector (Promega) (Supplementary Figure S4) were cloned upstream of the firefly luciferase sequences between the HindIII and KpnI restriction enzyme sites. The sequences of the cloned elements were verified (Eurofins). The viral BHLF1 promoter with the ZREs mutated was cloned upstream from the luciferase reporter gene in pCpGL luciferase reporter vector (34) and the Zta binding sites from the cellular FOSB gene and another region of cellular DNA were cloned upstream from it (Supplementary Figure S4). Zta binding sites ≥40kb downstream from the TSS of RASA3 were cloned into pGL3-enhancer reporter vector (Promega) and the minimal promoter from the pGL4.23 reporter vector was added (Promega) (Supplementary Figure S4).

Different DNA combinations comprising 5 µg of the relevant reporter construct and 5 µg of an expression vector for His-tagged Zta, or the corresponding empty vector control (35), were introduced into  $1.2 \times 10^7$  DG75 cells, an EBV-negative BL line (31), by electroporation with the Gene pulser II (Biorad). After 24 h, half of the cells were harvested into 250 µl of 1× passive lysis buffer (Promega), and luciferase activity was determined using the luciferase assay system (Promega). Light emission was measured in relative light units (RLU) using the GloMax multi detection system (Promega). The remaining cells were harvested into

**Table 1.** Primers used for Q-PCR assays

	5' flank	Zta peak	3' flank
<i>SCIMP</i>		Forward primer CCCTCGTGCAATACTGTGAGA Reverse primer ACAACTCATTCGCTCTGGGC	Forward primer TTGCACAGCAAGTTCAAGCC Reverse primer CTTTCTTGAAGG CAGATGGCAA
<i>BCL2A1</i>	Forward primer AGGAATTTGGCCTCCCAATCA Reverse primer TTTCTCCAGCGACCATGAGTT	Forward primer TCTTGAGCTGGCTCACCTTG Reverse primer AAACACAGCCTACGCACGAA	Forward primer ACAGTGGTTACC TCTTGGGAGA Reverse primer CCTGTGTTGAAA CTCATGTTGGTA



**Figure 1.** Zta binding across the human genome. ChIP-Seq was undertaken with an antibody against Zta protein using chromatin from Akata BL cells induced to enter EBV lytic replication cycle by cross-linking the B-cell receptor with anti-IgG. The DNA bound by Zta was aligned to the human genome and peaks of Zta binding were identified using model-based analysis of ChIP-seq (MACS). (A) The proportions of genes associated with proximal (<2 kb from TSS), intermediate (2–4 kb) and distal (>4 kb) Zta binding sites is shown. (B) The presence of common DNA sequence motifs within the Zta binding sites was determined using MEME-ChIP. The enriched motifs are colored together with the significance of the enrichment. The related ZRE motif is shown below each in black (17).

protein sample buffer and the levels of Zta were determined by immunoblotting (as described above). For each sample, relative promoter activity was determined from the ratio of luciferase activity to protein level and the assays were performed in triplicate.

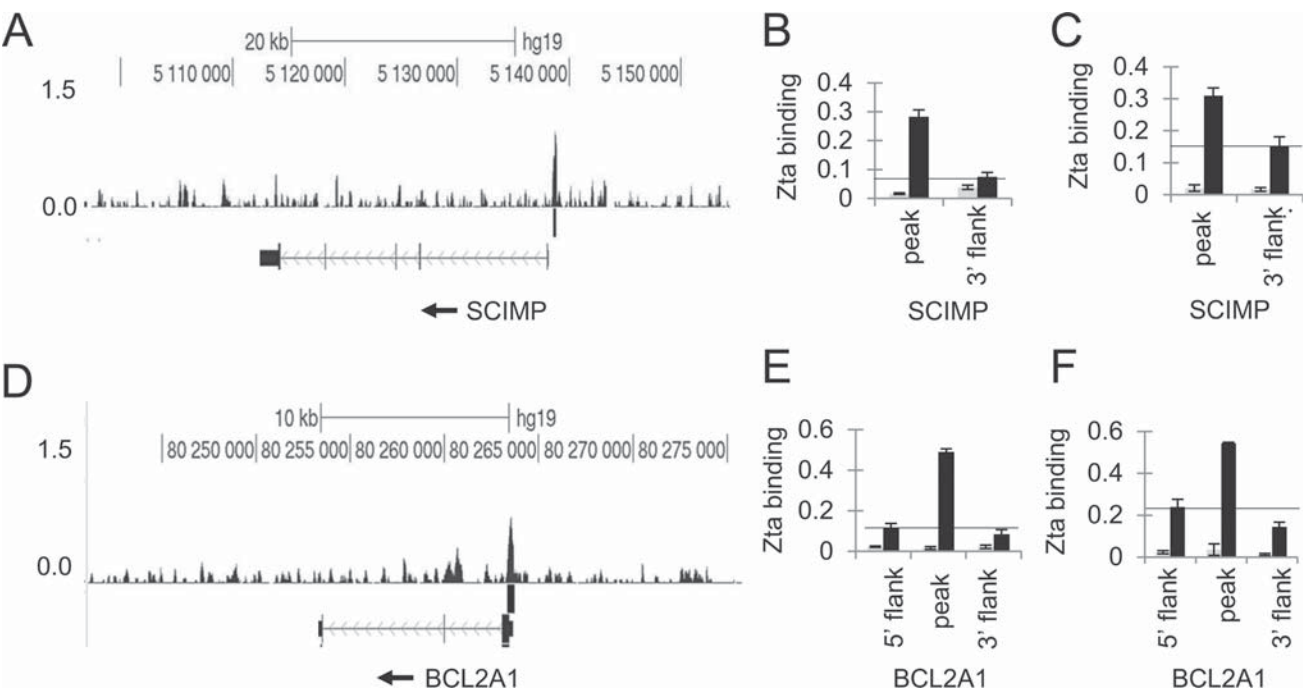
### ChIP

Surface IgG on Akata cells was cross-linked by application of anti-IgG and 48 h later chromatin was prepared from  $1 \times 10^8$  cells. An additional sample was treated with acyclovir to halt viral replication prior to EBV genome replication. The chromatin was fixed by application of 1% (v/v) formaldehyde for 15 min at 20°C, extracted as described in Bark-Jones *et al.* (36) then sonicated on ice ( $10 \times 10$  s-pulses; 30% amplitude output) on a Branson model 250 Microtip at setting 5 (Sonics Vibacell). The chromatin was pre-cleared with protein A/G-sepharose bead slurry (Sigma) that had been pre-blocked in 0.5% (w/v) fraction V BSA (Sigma) in DPBS. 2% (v/v) of the pre-cleared extract was retained as the input control sample while the remainder was incubated with 10 µg of Zta-specific goat antibody (Santa Cruz # sc-17503) or control goat IgG for 1 h at 4°C as described in Ramasubramanian *et al.* (20). The input control sample and the precipitated DNA were sequentially treated with 0.2 µg/ml RNaseA and 0.2 µg/ml Proteinase K and the DNA purified. Sequencing libraries were prepared using 10 ng of the input and ChIP DNA using a ChIP-seq sample preparation kit (Illumina) following the manufacturer's protocol, except that the library (150–350 bp fragments) was purified from the gel using a gel extraction kit (Qiagen) after PCR-amplification. The libraries were sequenced (single-

end, 36 bp) using an Illumina Genome Analyzer IIx. Initial processing of sequencing images was carried out using the CASAVA pipeline. Base calling, and quality control statistics were performed using GOAT and Bustard modules. Sequence reads were aligned to the hg19 release of the human genome with ELAND. Bigwig files were generated by calculating tag density in 10-bp window, normalizing per million total reads and subtracting input signals using in-house R scripts. Peak calling ( $P < 10^{-7}$ ) was performed from a merged Bed file generated from the two ChIP-datasets using MACS (37), with a merged input Bed file acting as background (GSE57732). Distances between binding sites and the transcription start site annotated in RefSeq (38) were calculated. Genes within 2 kb of a peak were scored as promoter proximal. The Zta binding sequences associated with peaks were analyzed for common DNA-sequence motifs using MEME-ChIP (39,40). Matches with both  $P$  and  $q$  values of  $<1.0E-034$  were selected and common microsatellite repeat sequences omitted. Functional gene enrichment analysis was undertaken using the DAVID bioinformatics resource at the National Institute of Allergy and Infectious Diseases (NIAID) (41,42).

ChIP was undertaken from chromatin generated from the spontaneously lytic LCL cell line LCL#3. Enrichment was quantified using Q-PCR spanning Zta binding peaks identified in the Akata cell line and flanking regions ~2 kb 5' or 3' to them. The DNA sequences of each of the primers is shown in Table 1.





**Figure 2.** Comparison of Zta binding patterns in BL, LCL and NPC cells. (A) Examples of ChIP-seq data from Akata cells aligned to the human genome surrounding the *SCIMP* locus (chr17:5 096 004–5 154 368, hg19) is shown using the UCSC genome browser. The depth of sequencing reads from Zta-enriched DNA is plotted per million background-subtracted total reads. The scale bar and genome location are shown, together with the Zta binding profile. Beneath this are the peaks of Zta binding identified using MACs ( $P < 10^{-7}$ ), with the RefSeq gene mapping shown below that. (B) The presence of Zta binding in LCL#3 is shown at the *SCIMP* locus using Q-PCR with primers corresponding to the Zta binding sites (peak) and a flanking region 2 kb distant (flank). The data from triplicate analyses show the percentage of binding relative to input chromatin. The DNA binding associated with Zta is shown in black and the control antibody in gray. A horizontal line marks the baseline for the assay. (C) The presence of Zta binding in HONE1-EBV at the *SCIMP* locus is shown (as in B). (D) Examples of ChIP-seq data from Akata cells aligned to the human genome surrounding the *BCL2A1* locus (chr15:80 241 085–80 275 791, hg19) is shown using the UCSC genome browser as in (A). (E) The presence of Zta binding in LCL#3 is shown at the *BCL2A1* locus using Q-PCR with primers corresponding to the binding sites and to two flanking regions 2 kb distant either side of the site as in (B). (F) The presence of Zta binding in HONE1-EBV cells at the *BCL2A1* locus is shown.

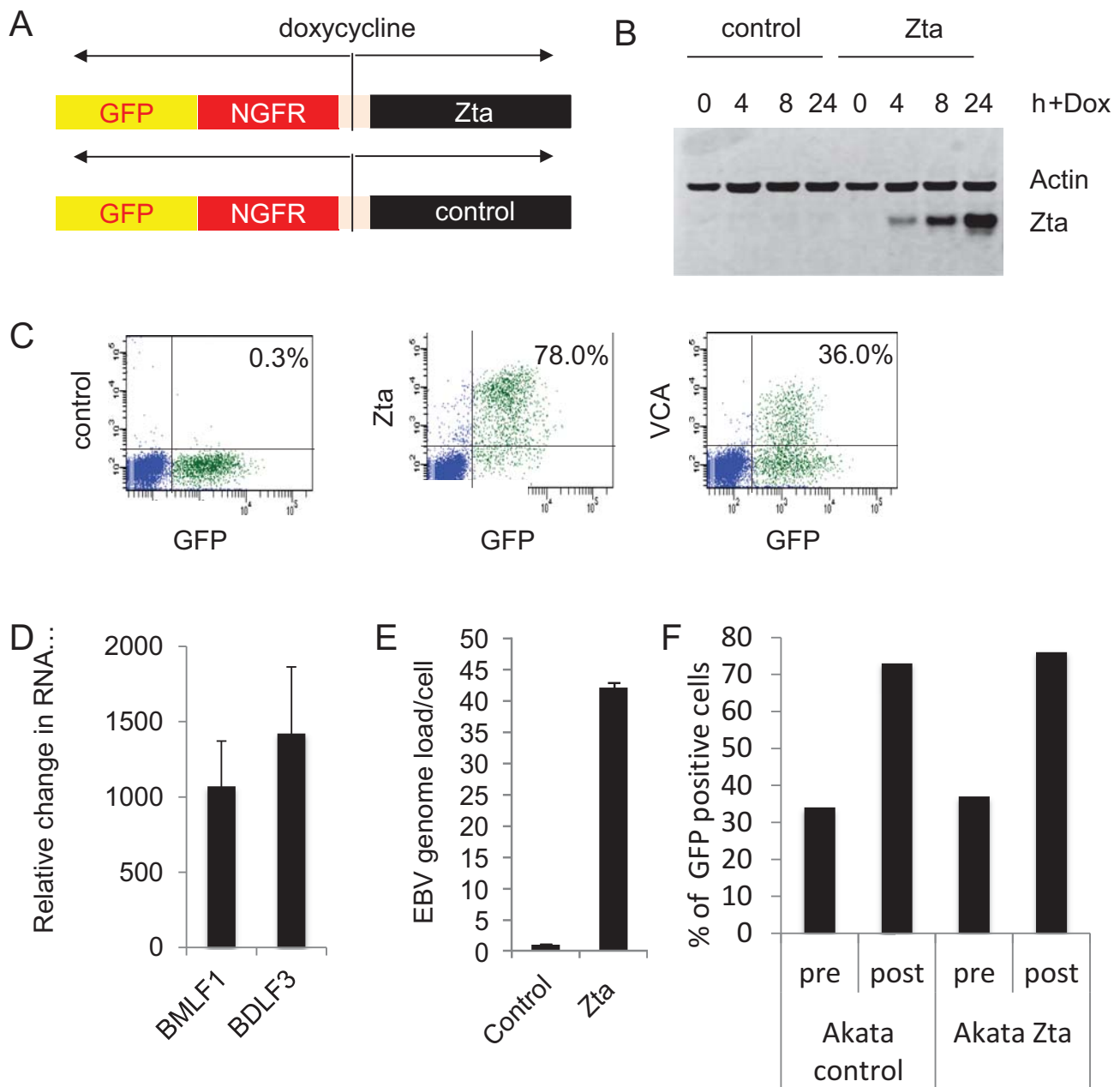
**Table 2.** Ten most enriched Gene Ontology biological process terms in the set of Zta-up regulated genes

GO term	Accession number	P Value	Bonferroni	Benjamini	FDR
Cell projection organization	GO:0030030	3.84E–14	1.34E–10	1.34E–10	7.04E–11
Cell morphogenesis	GO:0000902	5.75E–14	2.01E–10	1.01E–10	1.05E–10
Cellular component morphogenesis	GO:0032989	1.03E–13	3.62E–10	1.21E–10	1.89E–10
Biological adhesion	GO:0022610	1.08E–13	3.79E–10	9.48E–11	1.98E–10
Neuron projection development	GO:0031175	1.62E–13	5.67E–10	1.13E–10	2.97E–10
Cell adhesion	GO:0007155	2.27E–13	7.95E–10	1.33E–10	4.16E–10
Cell morphogenesis involved in differentiation	GO:0000904	7.67E–13	2.69E–09	3.84E–10	1.40E–09
Neuron projection morphogenesis	GO:0048812	1.28E–12	4.49E–09	5.62E–10	2.35E–09
Cell projection morphogenesis	GO:0048858	3.15E–12	1.10E–08	1.23E–09	5.77E–09
Cell morphogenesis involved in neuron differentiation	GO:0048667	7.83E–12	2.74E–08	2.74E–09	1.43E–08

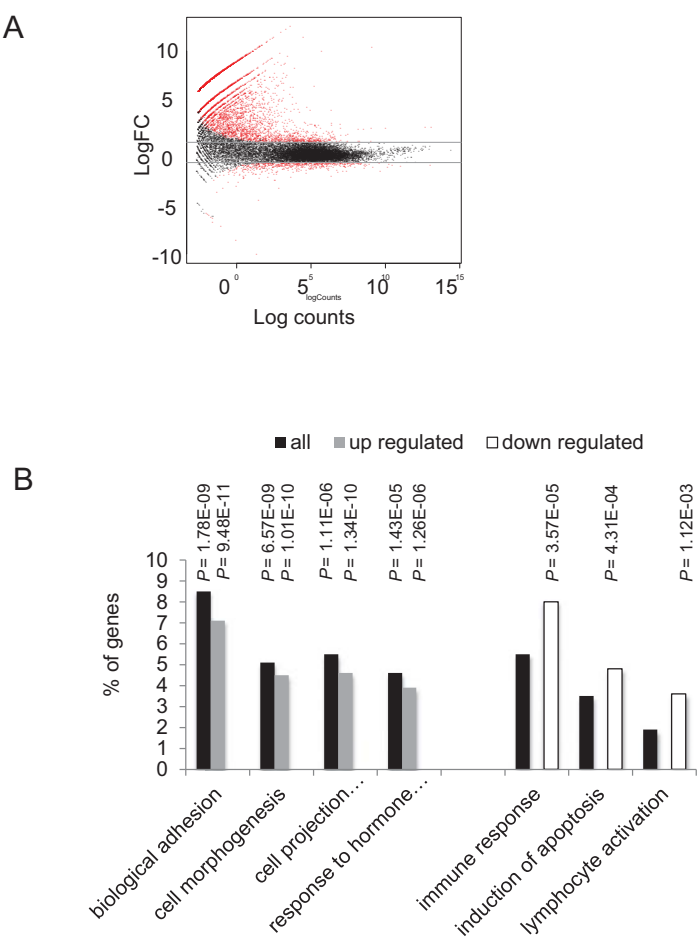
Reverse transcription and qPCR

Total RNA was prepared from Akata-Zta and Akata control cells 24 h after addition of doxycycline using RNAeasy Kit (Qiagen) and treated with RQ-DNase (Promega) and quantitated using a Nanodrop spectrometer. First strand cDNA was prepared using random primers (Roche). Viral transcripts were amplified using absolute quantitation with Sybrgreen Go-Taq Q-PCR (Promega) with the following primers: BMLF1 (GACCGCTTCGAG

TTCCAGAAT; ACTCTCCCGAACTAGCAGCAT and BDLF3 (TCGGTGGCAGTGATGTTCTG; TTCCAACGCATCCACCATCA). Host cell RNAs were also analyzed using TaqMan relative quantitation RT-PCR in a high-throughput format using Gene cards (Life Technologies, UK). Signals were normalized to GAPDH expression and the relative difference in the presence and absence of doxycycline determined. The fold change in RNA abundance was calculated together with the



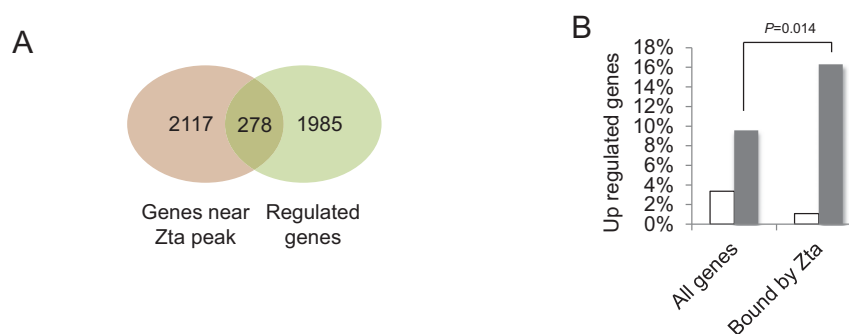
**Figure 3.** Lytic cycle induction system. (A) A schematic diagram of the bi-directional expression vector encoding Zta and is shown. The control vector has the Zta sequence cloned in the opposite direction. (B) Protein expression was assessed over 4, 8 and 24 h using immunoblotting for Zta and actin as indicated. (C) The expression of Zta and the late lytic cycle protein VCA were determined in Akata Zta cells induced with doxycycline for 24 h using multi-parameter FACS analysis. The data are expressed as the % of the GFP positive cells that also stain for Zta and VCA compared to a control antibody. (D) Expression levels of the viral genes BMLF1 and BDLF3 mRNA were assessed using RT-coupled Q-PCR. The relative change in abundance of the mRNA is shown with the error bars relating to triplicate cell induction experiments and standard deviation. (E) The relative amount of viral compared to human DNA (viral genome load) was determined by Q-PCR 24 h after doxycycline addition. This is shown with the mean and standard deviation from triplicate assays. (F) Cells were induced with doxycycline for 24 h and Zta expressing and control cells were enriched with anti-NGFR coated magnetic beads. The proportion of cells that express GFP is shown for both the pre-enriched and post-enriched populations of cells.



**Figure 4.** Global transcriptome analysis of gene expression changes reprogrammed during viral replication. (A) The expression of cellular genes in two enriched populations of Akata-control and Akata-Zta cells induced with doxycycline for 24 h were assessed using RNA-seq. The distribution of genes with a change in expression of at least two and a false discovery rate (FDR) ≤ 0.01 are shown in red on the MA plot (log total counts versus log fold-change). (B) Percentage of genes enriched (Bonferroni-modified  $P \leq 0.01$ ) for Biological Function categories. Data are shown for all genes (black), those that are up regulated (gray) and those that are down regulated (open).

**Table 3.** Ten most enriched Gene Ontology biological process terms in the set of Zta-down regulated genes

GO term	Accession number	P Value	Bonferroni	Benjamini	FDR
Immune response	GO:0006955	1.67E-08	3.57E-05	3.57E-05	2.89E-05
Induction of apoptosis	GO:0006917	2.02E-07	4.31E-04	2.16E-04	3.49E-04
Induction of programmed cell death	GO:0012502	2.15E-07	4.59E-04	1.53E-04	3.72E-04
Positive regulation of apoptosis	GO:0043065	2.66E-07	5.68E-04	1.42E-04	4.60E-04
Positive regulation of programmed cell death	GO:0043068	3.06E-07	6.55E-04	1.31E-04	5.30E-04
Positive regulation of cell death	GO:0010942	3.43E-07	7.35E-04	1.23E-04	5.95E-04
Lymphocyte activation	GO:0046649	5.24E-07	1.12E-03	1.60E-04	9.08E-04
Cell activation	GO:0001775	1.14E-06	2.45E-03	3.06E-04	1.98E-03
Regulation of apoptosis	GO:0042981	3.37E-06	7.18E-03	8.01E-04	5.84E-03
Regulation of programmed cell death	GO:0043067	4.38E-06	9.33E-03	9.37E-04	7.59E-03



**Figure 5.** Intersection of ChIP-seq and RNA-seq data. (A) The diagram shows the intersection between genes associated with a Zta binding peak (brown) and genes that are regulated during EBV lytic cycle (green). (B) The percentage of all cellular genes and the percentage of the sub-set of cellular genes that are associated with Zta binding and up regulated during lytic cycle was compared. The significant increase in the percentage of up regulated genes that are also associated with Zta binding is highlighted ( $P < 0.01$ , Binomial distribution (cumulative),  $n = 92$ ).

standard deviation from three independent cell enrichment experiments. The TaqMan primer sets were as follows: GAPDH-Hs99999905\_m1; FOSB-Hs00171851\_m1; SCIMP-Hs010294\_m1; AN01-Hs00216121\_m1; GDF2-Hs00211913\_m1; FSCN1-Hs00602051\_mH (Life Technologies, UK).

### RNA-Seq

Sequencing was performed on biological duplicates of Akata-Zta and Akata control samples on the Illumina HiSeq 2500 platform and generated ~45 million 101-bp paired-end reads per sample. The data are deposited in GEO (GSE57732). Sequenced reads were mapped to RefSeq genes archived in the Illumina iGenomes resource ([https://support.illumina.com/sequencing/sequencing\\_software/igenome.ilm](https://support.illumina.com/sequencing/sequencing_software/igenome.ilm)) using RSEM (version 1.2.4; (43)). An existing pipeline developed within the Trinity software package ([http://trinityrnaseq.sourceforge.net/analysis/diff\\_expression\\_analysis.html](http://trinityrnaseq.sourceforge.net/analysis/diff_expression_analysis.html)) (44) was used to perform differential expression analysis with edgeR (version 3.01; <http://www.bioconductor.org/packages/release/bioc/html/edgeR.html>), which is available as part of the Bioconductor project developed within the R programming language (45). Genes with log CPM > 0, fold-changes above 2-fold and FDR < 0.01 were judged to be differentially expressed. Genes with a total read count below 20 were excluded from further analysis. Views of the data are plotted using the Integrative Genomics Viewer (IGV) (46).

### DNA analysis

Total cell DNA was prepared from cells and the amounts of EBV genome relative to human genome determined using Q-PCR as described (47).

## RESULTS

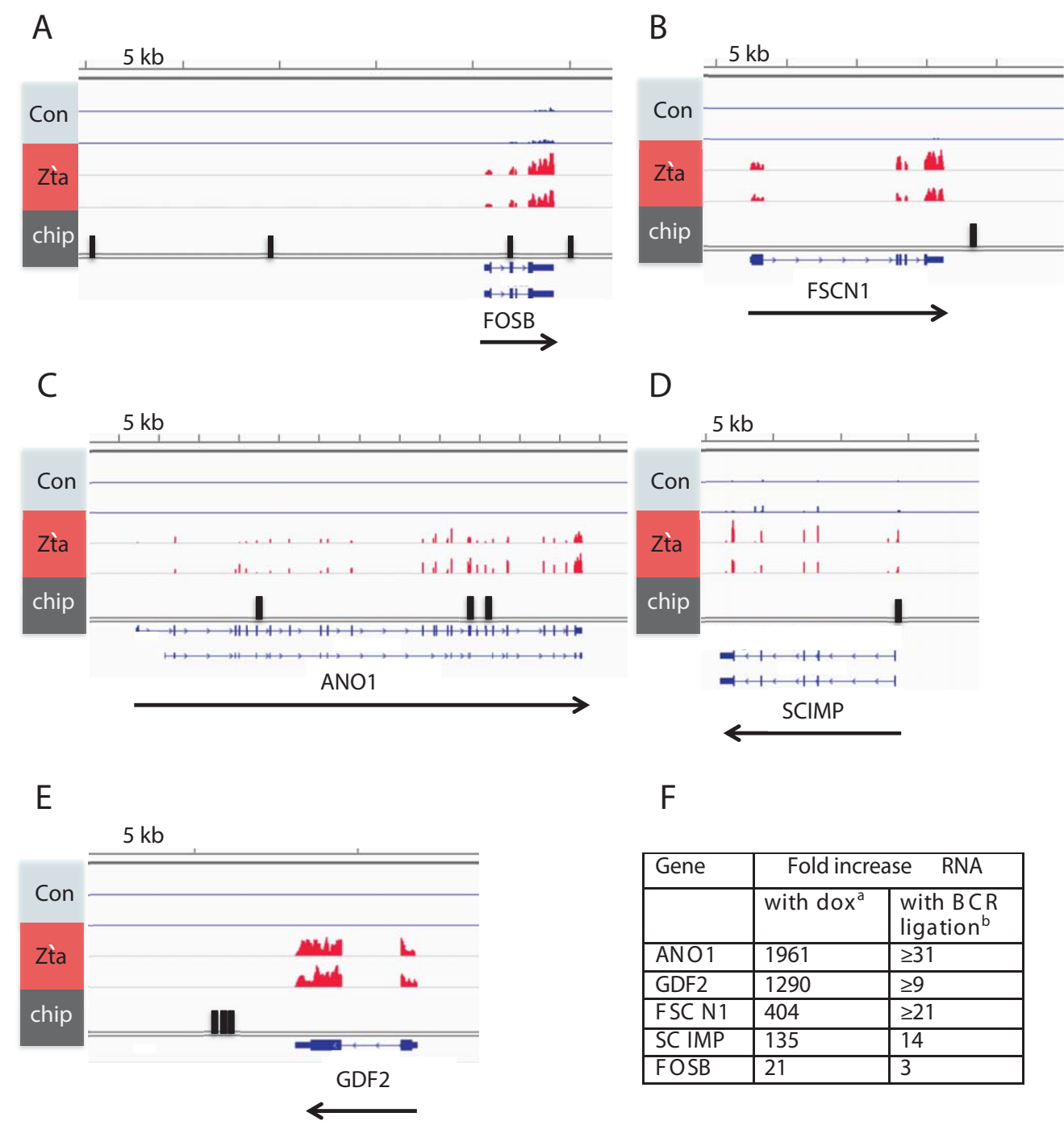
### Candidate targets of Zta in the human genome

To determine whether Zta interacts with human genes during the switch from latency to lytic EBV replication, we undertook a genomic-scale analysis of Zta binding sites in

Akata EBV-positive BL cells. Lytic replication was activated by engagement of the cell surface B-cell receptor (BCR) with anti-immunoglobulin. After crosslinking the proteins to DNA using formaldehyde, chromatin from the activated cells was immunoprecipitated with a previously validated Zta antibody (20,47) and the precipitated DNA was analyzed by massively parallel DNA sequencing (ChIP-seq). Significant regions of interaction between Zta and the cellular genome were identified using the Model-based Analysis of ChIP-Seq (MACS) algorithm (37). This identified 5020 Zta binding peaks ( $P \leq 1 \times 10^{-7}$ ). Using a motif-based sequence analysis tool (MEME-ChIP) (39,40) to search for enriched motifs we noted that the most frequent consensus sequences resembled the previously described non-CpG containing and CpG containing ZRE motifs from the EBV genome (16,17,20), suggesting that Zta directly binds to these sites in the human genome.

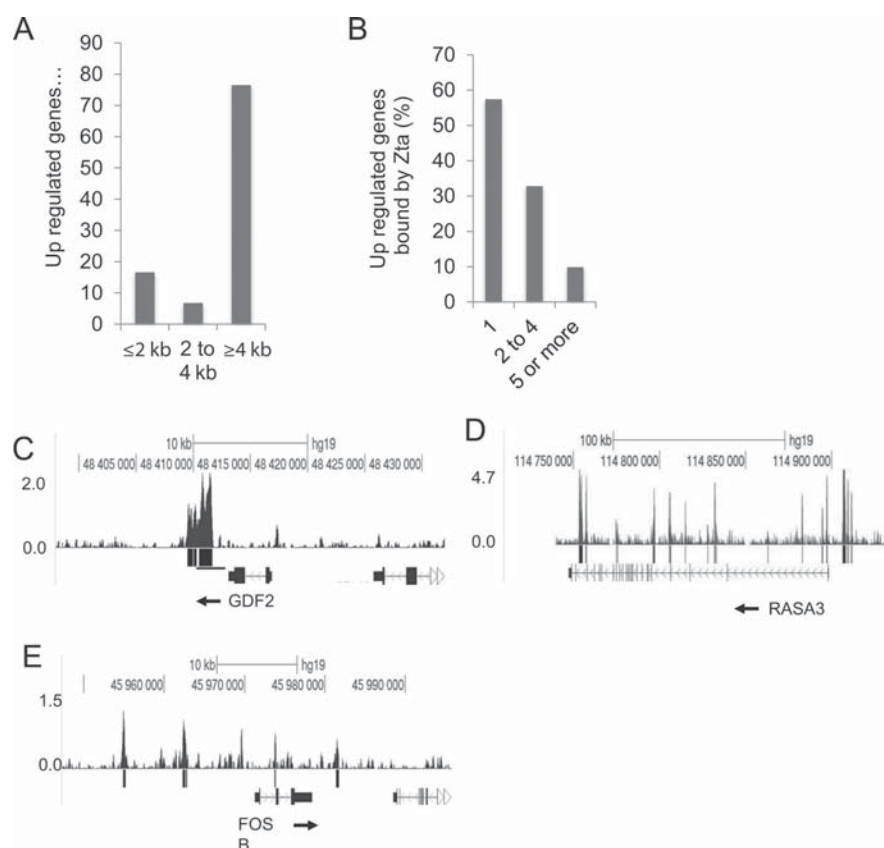
Although some of the Zta binding sites (14%) were close to the annotated transcriptional start sites (TSS) of cellular genes (Figure 1B and Supplementary Table 1), analogous to the situation observed for the interaction of Zta with viral genes, the majority of the Zta binding sites (86%) were more distant ( $\geq 4$  kb from the nearest TSS). By assigning each of the Zta peaks to the closest TSS, irrespective of distance, we identified 2395 cellular genes as potential candidates for regulation by Zta.

To validate the ChIP-seq, we chose representative examples of cellular genes associated with Zta binding sites, based on the patterns of Zta binding and amenability to analysis by Q-PCR using PCR primer sets that would discriminate enhanced binding within the peak region relative to the adjacent DNA. We also determined whether Zta binding occurs in different cell types in which EBV undergoes lytic replication: the LCL#3 lymphoblastoid cell line (30) and HONE1-EBV, a cell line derived from a nasopharyngeal carcinoma (48). Examples of the analysis of Zta binding at the *SCIMP* and *BCL2A1* genes are shown in Figure 2A and D. In both cases, Zta binding was enriched within the peak region identified by ChIP-seq in Akata cells undergoing lytic replication (Figure 2B, C, E and F). This demonstrates that the specific interaction between Zta and these genes occurs in three different cell backgrounds.



**Figure 6.** Changes in RNA abundance for five cellular genes after induction of EBV lytic replication cycle in Akata cells. (A–E) The raw RNA-seq reads from two independent isolations of mRNA for control and Zta expressing cells are aligned to the human genome at *ANO1*; *GDF2*; *FSCN1*; *SCIMP* and *FOSB* loci using the IGV genome browser. A scale bar is shown above each region. The location of the Zta binding sites is shown in black. (F) RNA was prepared from Akata cells following crosslinking BCR with anti-IgG for 48 h. Changes in the expression of each gene: *ANO1*; *GDF2*; *FSCN1*; *SCIMP* and *FOSB* were determined using TaqMan RT-PCR. The fold induction data is shown (with BCR ligation, b), together with the fold change from the doxycycline induced Akata Zta and Akata controls cells (with dox, a).





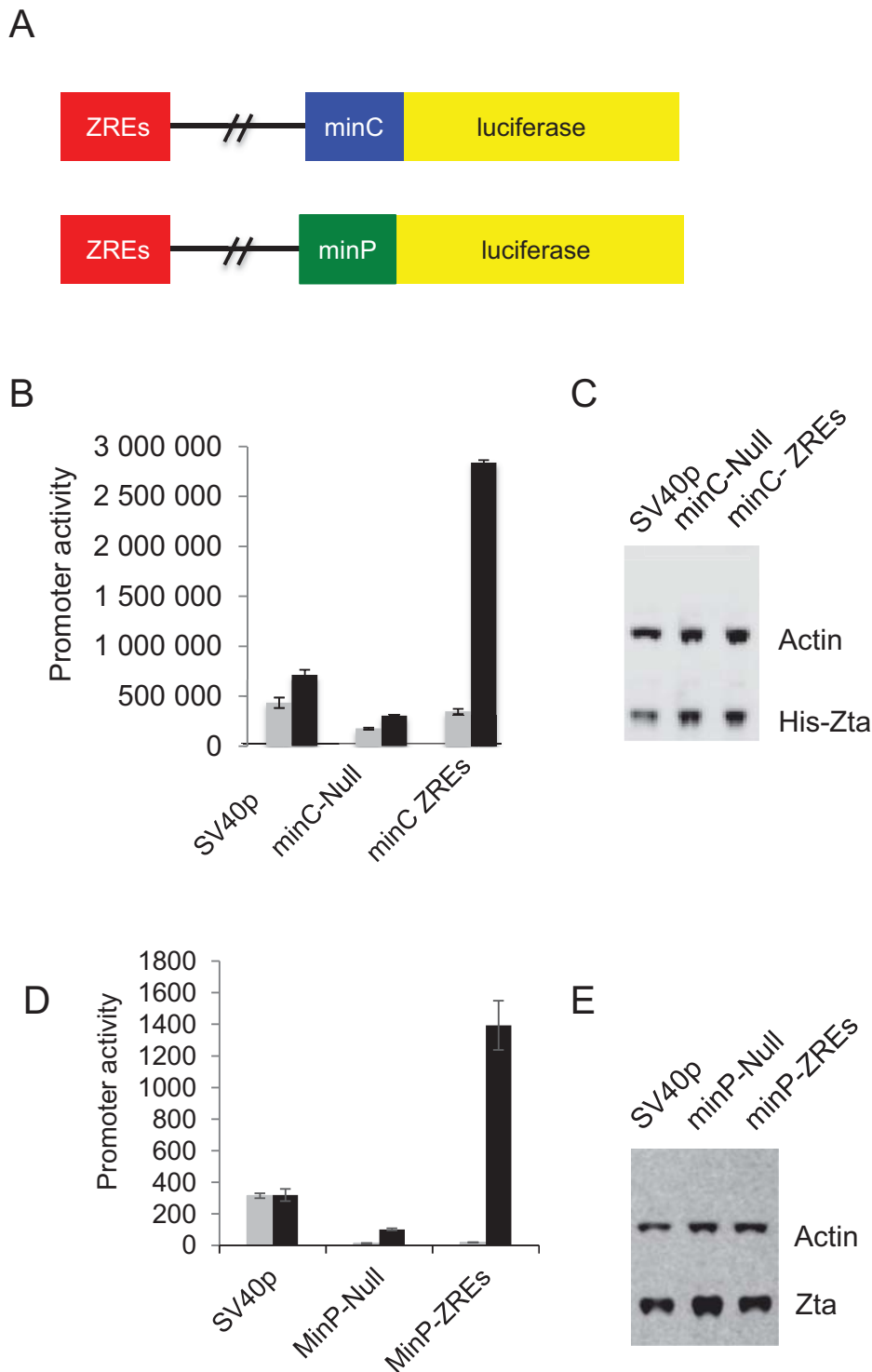
**Figure 7.** Location of Zta binding sites at up regulated genes. (A) The figure shows the percentage of up regulated genes with Zta binding sites within 2 kb, between 2 and 4 kb and >4 kb from a TSS. (B) The percentage of up regulated genes with 1, 2–4 or >5 Zta binding sites. (C) The numbers of sequencing reads from Zta-enriched DNA at the *GDF2* locus (chr10:48 397 968–48 431 996, hg19) are plotted per million background-subtracted total reads and aligned with the human genome. The scale bar and genome location are shown, together with the Zta binding. Beneath this are the peaks identified using MACS ( $P < 10^{-7}$ ) and RefSeq gene mapping (UCSC genome browser). (D) Similar data for the region surrounding the *RASA3* locus (chr13:114 709 728–114 936 080, hg19). (E) Similar data for the region surrounding the *FOSB* locus (chr19:45 947 279–45 995 778 hg19).

### Changes in the cell transcriptome during EBV lytic cycle

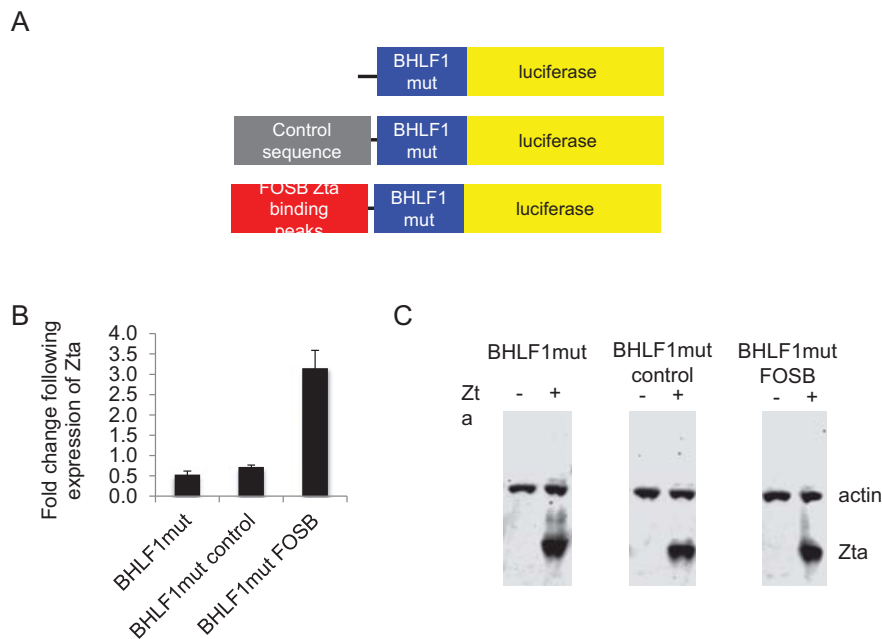
To explore the significance of Zta binding throughout the genome of Akata cells, we wanted to establish whether any of the genes associated with Zta binding sites are transcriptionally regulated during the EBV lytic replication cycle. Previous attempts to investigate changes in cellular gene expression during the EBV lytic cycle have been limited by the low percentage of Akata cells that undergo EBV lytic replication following BCR engagement. To optimize the identification of regulated genes, we introduced an inducible Zta expression vector with a selectable marker gene (32) into Akata cells to allow lytic cycle induction in a high proportion of cells (Figure 3A). In parallel, a cell line was generated with a vector in which the BZLF1 sequences were inserted in the opposite orientation. This cell line was used as a comparator throughout the investigation to control for any impacts of the inducing agent on cells. Addition of doxycycline resulted in the rapid induction of Zta expression (Figure 3B and C) with no associated toxicity (Supplementary Figure S1). This was accompanied by activation of the EBV lytic cycle, as judged by the expression of the late gene VCA (Figure 3C), expression of the viral early gene BMLF1 and late gene BDLF3 (Figure 3D) and replication of the viral

genome (Figure 3E). No evidence for lytic replication was observed in the control cells. It is important to note that although a higher percentage of cells express Zta protein in the inducible system compared to BCR stimulation, the abundance of Zta protein is equivalent in the two systems as judged by intracellular staining and FACS analyses (Supplementary Figure S2). Thus, the level of Zta expression induced by doxycycline is physiologically relevant.

Populations of control and Zta expressing cells were enriched by selection on anti-NGFR coupled magnetic beads. This generated pools of cells in which greater than 70% expressed the induced genes (Figure 3F and Supplementary Figure S3), in which both the down regulation and the up regulation of gene expression can be determined. This was undertaken for two independently enriched populations of Zta expressing and control cells by sequencing the polyadenylated transcripts (RNA-seq). Differential expression analysis identified 2263 cellular genes whose expression changed >2-fold following induction of the lytic cycle ( $\text{FDR} \leq 0.01$ ) (Figure 4 and Supplementary Table S2). This included 2242 novel targets and 21 genes that had been identified in previous studies (27,28). Of the genes identified in this analysis, 1679 were up regulated and 584 were down regulated. The genes fell into many functional groups



**Figure 8.** Zta can activate gene expression through a long-range element. (A) A synthetic element containing tandem ZREs was cloned 2.2 kb from either of two minimal promoters (minC and minP) in the pGL3 control luciferase reporter plasmid. (B) Promoter activities of the pGL3 based plasmids SV40p, minC-Null and minC ZREs in the presence (black) and absence (gray) of His-Zta expression in the EBV negative BL cell line (DG75). The relative luciferase activity is shown as promoter activity together with the standard deviation from three replicates. (C) Western blot analysis of Zta and Actin from samples in (B). (D) Promoter activity of the minP-Null and minP-ZREs reporters in the presence and absence of His-Zta expression in the EBV negative BL cell line (DG75). The relative luciferase activity is shown as promoter activity together with the standard deviation from three replicates. (E) Western blot analysis of Zta and actin from samples in (B).



**Figure 9.** Zta can activate gene expression through binding sites associated with *FOSB*. (A) The DNA sequences at three Zta binding sites associated with *FOSB* were cloned upstream from an EBV lytic cycle promoter BHLF1 containing no ZREs (BHLF1mut) to generate BHLF1mut FOSB. A control region from the human genome is also included. (B) Promoter activity of BHLF1mut, BHLF1mut control and BHLF1mut FOSB in the presence and absence of His-Zta expression in the EBV negative cell line (DG75). The relative luciferase activity is shown as fold regulation by Zta with the standard deviation from three replicates. (C) Western blot analysis of Zta and Actin from samples in (B).

but gene ontology analyses showed that among the up regulated genes, significant enrichment was observed for genes involved in cell adhesion, morphogenesis, projection and response to hormones (Table 2 and Figure 4). Within the down regulated category, there was significant enrichment for genes involved in the immune response, induction of apoptosis and lymphocyte activation (Table 3 and Figure 4).

### Regulation of Zta-associated genes

Although the presence of a transcription factor on a gene can be indicative of a regulatory role, it is not definitive: transcription factor binding can occur without discernible changes in gene expression. We therefore asked which of the 2395 cellular genes associated with Zta binding are regulated following Zta induction of the EBV lytic cycle. Of the genes whose expression changed  $\geq 2$ -fold in this experiment, 278 were associated with Zta binding sites and are therefore candidates for direct Zta-mediated regulation (Figure 5A). Of these, 207 genes were up regulated and 71 were down regulated. The presence of one or more Zta binding site correlated with an increased likelihood that a gene was up regulated ( $P = 0.014$ , Binomial distribution,  $n = 92$ ) (Figure 5B). In contrast, there was no specific enrichment for Zta binding sites in the down regulated genes.

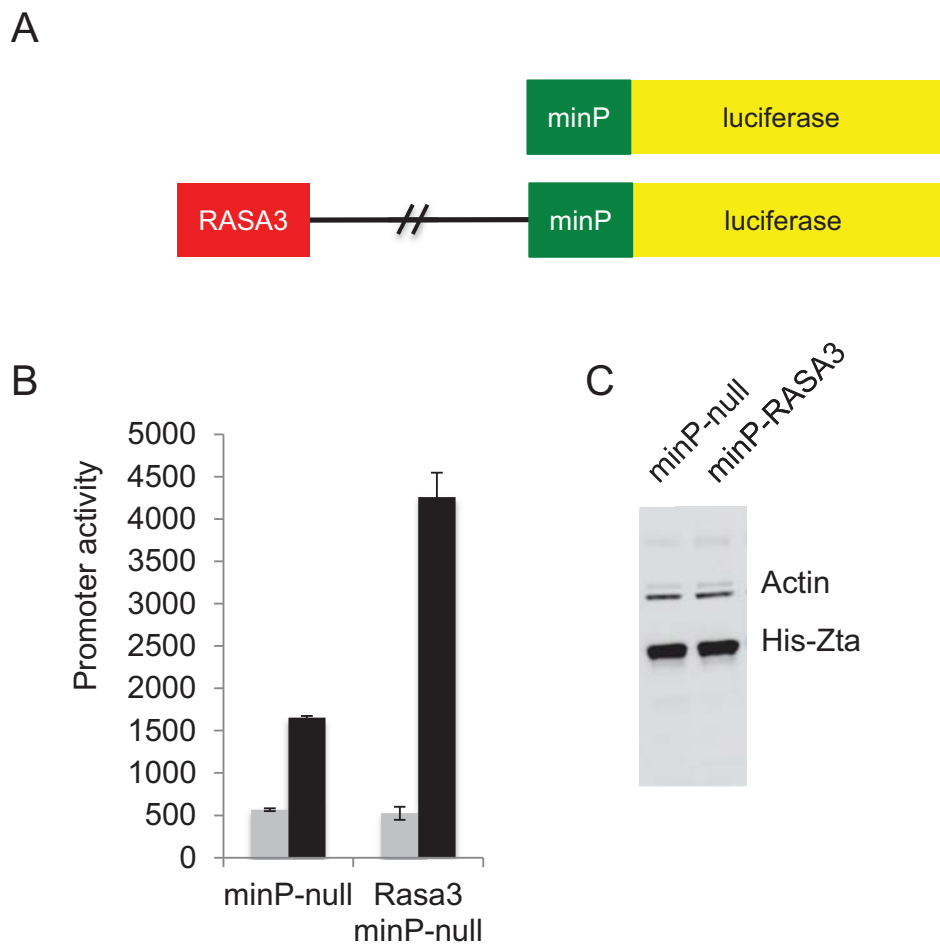
The proportion of genes associated with Zta binding that are demonstrably regulated during EBV lytic replication in BL cells (11.6%) is typical of many cellular transcription factors. Published reports suggest that between 1% and 10% of transcription factor associated genes are transcriptionally altered by expression of the transcription factor (49–52).

The Zta-bound genes that are also up regulated during EBV lytic cycle, listed in Supplementary Table 3, are the most likely candidates to be regulated directly by Zta and we focused our follow-up analysis on these. We chose a panel of five representative genes in which the Zta binding sites were in different positions relative to the TSS: *SCIMP1*, *FOSB*, *ANO1*, *GDF2* and *FSCN11*. Figure 6 shows the aligned RNA sequence data and the locations of Zta binding sites. For each gene, we obtained appropriate Taq-man probes and quantified the changes in abundance of the RNA in Akata cells following the ligation of BCR by cross-linking with anti-IgG. All of the RNAs showed increased abundance following BCR ligation (Figure 6E), although the effects were difficult to quantify because signals in the unligated cells were below the levels of detection.

### Patterns of Zta binding at regulated genes

Having validated the RNA-seq data at selected targets, we asked whether there was any general correlation between the number and location of Zta binding sites and the observed effects on gene expression at up regulated genes (Figure 7). Although  $\sim 15\%$  of the Zta binding sites were within 2 kb of a TSS, in line with the established model of Zta mediated gene activation of viral promoters, it was clear that the majority (75%) of the Zta binding sites were distal ( $>4$  kb) to the nearest TSS. Approximately half of the up regulated genes associated with a Zta binding site have a single peak whereas the others have two or more (Figure 7).

At some genes, Zta binding was relatively confined as exemplified by the up regulated *GDF2* gene where a cluster of four Zta binding sites was observed at the 3' end of the gene,



**Figure 10.** Zta can activate gene expression through binding sites associated with *RASA3*. (A) The DNA sequences at seven Zta binding sites associated with the 3' region of *RASA3* were cloned in the enhancer position >2 kb from the minP promoter to generate minP-RASA3. (B) Promoter activity of minP null and RASA3 minP in the presence and absence of His-Zta expression in the EBV negative cell line (DG75). The relative luciferase activity is shown with the standard deviation from three replicates. (C) Western blot analysis of Zta and Actin from samples in (D).

6–8 kb from the transcription start site (Figure 7C). However, at other genes, exemplified by *RASA3* and *FOSB*, Zta binding peaks were found 5' of the transcription start sites, within intragenic regions and 3' of the genes, distributed over considerable distances (Figure 7D and E).

**Zta binding sites act as transferable regulatory elements**

The ability of Zta to activate transcription by binding to promoter elements is well documented (53) but we are not aware of any previous reports that Zta can act through distal regulatory elements. To gain support for the idea that Zta can also influence transcription from a distance, we generated luciferase reporter constructs in which a tandem array of non CpG ZREs was cloned into a classical enhancer vector, 2.2 kb upstream of one of two minimal promoters (Figure 8A and Supplementary Figure S2). Tandem arrays of ZREs embedded within minimal promoters have been previously used to measure Zta activity (54,55), but action away from a promoter has not been assessed. In the MinC-ZREs reporter, we used a minimal promoter derived from

the *CIITA* gene, which is not regulated by Zta. An alternative reporter, designated MinP-ZREs, was based on the minimal promoter in the commercially sourced pGL4.23 reporter vector (Promega). In each case, we compared the activity of the promoters with and without the ZREs, with an SV40 promoter construct as an additional control (Figure 8B and D). The reporter plasmids were introduced into EBV negative DG75 BL cells, together with a vector encoding Zta or empty vector control. Immunoblotting confirmed that Zta was expressed at equivalent levels in the transfected cell populations (Figure 8C and E). Whereas Zta had a minimal effect on the activity of the MinC and SV40 promoters, the inclusion of the ZRE array resulted in an 8-fold increase in the activity of the MinC construct. Similarly, although Zta appeared able to activate the MinP promoter on its own, the presence of the ZRE array resulted in more substantial activation (75-fold versus 13-fold). The 6-fold differential between these effects is consistent with the impact of the ZRE array on the MinP promoter. Taken together, the data suggest that Zta can regulate gene expres-

sion via ZRE elements that are located at a distance from promoters, in a classical enhancer assay.

We then asked whether authentic Zta binding sequences from the cellular targets genes can be transferred onto a heterologous promoter. We chose the *FOSB* and *RASA3* genes as examples. In *FOSB* each of the four non-CPG Zta binding sites was  $\geq 2$  kb from the *FOSB* TSS (16 and 9 kb upstream, 2.4 and 10 kb downstream) and the gene is up regulated during the EBV lytic replication cycle (Figure 7E). We cloned the Zta binding sites from  $-16$  and  $-9$  kb upstream from a viral promoter (BHLF1mut) that had been rendered unresponsive to Zta by virtue of mutations in each of its ZREs (Figure 9A). We also included a similar size control region from the cellular genome. The reporter plasmids were transfected into DG75 cells, together with a plasmid encoding Zta or the empty vector control. The presence of the *FOSB* sequences caused a 2.5-fold increase in the response of the promoter to Zta compared to the controls (Figure 9B). Equivalent expression of Zta protein was seen for all reporters (Figure 9C). We therefore conclude that the distal Zta binding sites from the *FOSB* gene are able to confer Zta responsiveness. For *RASA3*, where there are multiple Zta binding sites across a  $>100$  kb region, we focused on elements within seven Zta binding sites from the 3' end of the *RASA3* gene ( $\geq 40$  kb from the TSS). These conferred Zta-mediated induction of expression of the minP promoter when cloned into the enhancer position of the PGL3 luciferase reporter vector. The average increase in expression was 2.2-fold over three experiments (Figure 10), which is in line with typical ESC enhancers in similar assays (56).

## DISCUSSION

We identified 2263 cellular genes whose expression is significantly altered within the first 24 h of initiation of the EBV lytic cycle. This was surprising given that a process termed host cell shut-off, driven by the EBV lytic replication cycle gene BGLF5, is reported to promote a global reduction in cellular gene expression during the EBV lytic replication cycle (57–59). The regulation of cellular genes that we observe must therefore occur either in advance of the impact of BGLF5 gene expression or perhaps in spite of it.

It is possible that the cellular genes that are regulated during the EBV lytic replication cycle are specifically targeted to aid the success or the efficiency of viral replication. Indeed, we found a significant enrichment for genes involved in specific biological functions. For example, 190 of the up regulated genes are involved in cell morphogenesis ( $P \leq 1.0 \times 10^{-10}$ ). We speculate that these genes might facilitate viral assembly or genome encapsulation in the nucleus, the export of immature virions from the nucleus, transport and further assembly within the cytoplasm, or egress from the cell. Among the down regulated group of genes, there was a specific enrichment for genes involved in apoptosis. It is well established that BL cells undergoing EBV lytic replication are protected from apoptosis (60) and the down regulation of the cellular genes identified here might contribute to the observed protection. This suggests that EBV might have evolved a strategy to reprogram cellular gene expression in order to tailor the environment of the cell for optimal viral replication.

During the early stages of the EBV lytic replication cycle, several viral genes, including the transcription factors Zta and Rta and the post-transcriptional regulator BSLF2 + BMRF1, act in concert to regulate the expression of the viral genome (53). In this scenario, it is clear that induction of Zta expression in Akata cells could drive both direct regulation of the host cell transcriptome and instigate indirect regulation of genes through the action of other lytic cycle proteins. It was therefore important to consider which of the 2263 host genes whose expression was altered following Zta activation were associated with a Zta binding peak during the EBV lytic replication cycle. Integration of the datasets provided 278 cellular genes as candidates for direct Zta-mediated transcriptional regulation.

One of the unexpected findings was that the majority of the Zta binding sites identified in this set of cellular genes are distal to promoters. This is very different from the situation observed in the viral genome where functional ZREs have been shown to lie in very close proximity to transcription start sites (16,20). Indeed, the interaction of Zta with RNA polymerase II accessory proteins is considered to aid the activation of transcription from viral promoters (61–64). The location of the cellular Zta binding sites suggested to us that Zta reprograms the regulation of cellular genes through a different mode of action. In support of this idea, we present two lines of evidence. First, we find that a tandem array of ZREs has enhancer-like activity when located distal to a minimal promoter in a classical enhancer reporter assay. Second, we find that the Zta binding sites from the cellular *FOSB* and *RASA3* genes can transfer Zta-mediated activation onto heterologous promoters. Although the impact is modest, it is in line with the magnitude of regulation recently reported for cellular enhancers in the context of plasmid based reporter assays (56). As the reporter gene regulation occurs in the absence of other EBV gene expression, it supports the hypothesis that Zta is able to directly activate cellular genes through regulatory elements that lie distal to their transcription start sites. This raises the possibility that long-range gene mechanisms act through directing through chromatin looping, in a similar manner to that recently described for the EBV EBNA genes (65,66). In light of the discovery of Zta regulation through distal Zta binding sites, it will be relevant to reconsider the interactions of Zta with the viral genome (16,20), with regards to the potential for regulation of distant promoters.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

Sequencing of the ChIP-seq libraries was performed by UCL Genomics. We are grateful to Nik Matthews and the Advanced Sequencing Facility at CRUK-London Research Institute for library preparation and RNA sequencing.

## ACCESSION NUMBERS

ChIP-sequencing and RNA sequencing data has been deposited in GSE57732.



## FUNDING

Medical research Council UK [MR/J001708/1 to A.J.S. and M.R.]; MRC Career Development Award [G0802068 to R.G.J.]; MRC Centre for Medical Molecular Virology and the National Institute for Health Research University College London Hospitals Biomedical Research Centre (to R.G.J.); Core funding to the London Research Institute from Cancer Research UK (to H.P. and G.P.). Funding for open access charge: University of Sussex through MRC grant [MR/J001780/01].

*Conflict of interest statement.* None declared.

## REFERENCES

- Rowe, M., Kelly, G.L., Bell, A.I. and Rickinson, A.B. (2009) Burkitt's lymphoma: the Rosetta Stone deciphering Epstein-Barr virus biology. *Semin. Cancer Biol.*, **19**, 377–388.
- Molyneux, E.M., Rochford, R., Griffin, B., Newton, R., Jackson, G., Menon, G., Harrison, C.J., Israels, T. and Bailey, S. (2012) Burkitt's lymphoma. *Lancet*, **379**, 1234–1244.
- Kuppers, R. (2009) The biology of Hodgkin's lymphoma. *Nat. Rev. Cancer*, **9**, 15–27.
- Young, L.S. and Rickinson, A.B. (2004) Epstein-Barr virus: 40 years on. *Nat. Rev. Cancer*, **4**, 757–768.
- Vetsika, E.-K. and Callan, M. (2004) Infectious mononucleosis and Epstein-Barr virus. *Expert Rev. Mol. Med.*, **6**, 1–16.
- Longnecker, R., Kieff, E. and Cohen, J.I. (2013) In: Knipe, D.M. and Howley, P. (eds). *Fields Virology*. 6th edn. Lippincott Williams, Philadelphia.
- Rickinson, A. and Kieff, E. (eds). (2007) *Epstein-Barr Virus*. 5th edn. Lippincott Williams and Wilkins.
- Kenney, S.C. (2007) In: Aea, Arvin (ed). *Reactivation and Lytic Replication of EBV*. 2011/02/25 edn. Cambridge University Press, Cambridge.
- Miller, G., El-Guindy, A., Countryman, J., Ye, J. and Gradoville, L. (2007) Lytic cycle switches of oncogenic human gammaherpesviruses(1). *Adv. Cancer Res.*, **97**, 81–109.
- Kalla, M. and Hammerschmidt, W. (2012) Human B cells on their route to latent infection—early but transient expression of lytic genes of Epstein-Barr virus. *Eur. J. Cell Biol.*, **91**, 65–69.
- Kalla, M., Schmeink, A., Bergbauer, M., Pich, D. and Hammerschmidt, W. (2010) AP-1 homolog BZLF1 of Epstein-Barr virus has two essential functions dependent on the epigenetic state of the viral genome. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 850–855.
- Shannon-Lowe, C., Adland, E., Bell, A.I., Delecluse, H.J., Rickinson, A.B. and Rowe, M. (2009) Features distinguishing Epstein-Barr virus infections of epithelial cells and B cells: viral genome expression, genome maintenance, and genome amplification. *J. Virol.*, **83**, 7749–7760.
- Halder, S., Murakami, M., Verma, S.C., Kumar, P., Yi, F. and Robertson, E.S. (2009) Early events associated with infection of Epstein-Barr virus infection of primary B-cells. *PLoS One*, **4**, e7214.
- Ma, S.D., Yu, X., Mertz, J.E., Gumperz, J.E., Reinheim, E., Zhou, Y., Tang, W., Burlingham, W.J., Gulley, M.L. and Kenney, S.C. (2012) An Epstein-Barr Virus (EBV) mutant with enhanced BZLF1 expression causes lymphomas with abortive lytic EBV infection in a humanized mouse model. *J. Virol.*, **86**, 7976–7987.
- Israel, B.F. and Kenney, S.C. (2005) In: Robertson, E.S. (ed). *Epstein Barr Virus*. Caister Academic Press, Wymondham, pp. 571–611.
- Bergbauer, M., Kalla, M., Schmeink, A., Gobel, C., Rothbauer, U., Eck, S., Benet-Pages, A., Strom, T.M. and Hammerschmidt, W. (2010) CpG-methylation regulates a class of Epstein-Barr virus promoters. *PLoS Pathog.*, **6**, e1001114.
- Flower, K., Thomas, D., Heather, J., Ramasubramanian, S., Jones, S. and Sinclair, A.J. (2011) Epigenetic Control of Viral Life-Cycle by a DNA-Methylation Dependent Transcription Factor. *PLoS One*, **6**, e25922.
- Bhende, P.M., Seaman, W.T., Delecluse, H.J. and Kenney, S.C. (2004) The EBV lytic switch protein, Z, preferentially binds to and activates the methylated viral genome. *Nat. Genet.*, **36**, 1099–1104.
- Bhende, P.M., Seaman, W.T., Delecluse, H.J. and Kenney, S.C. (2005) BZLF1 activation of the methylated form of the BRLF1 immediate-early promoter is regulated by BZLF1 residue 186. *J. Virol.*, **79**, 7338–7348.
- Ramasubramanian, S., Kanhere, A., Osborn, K., Flower, K., Jenner, R.G. and Sinclair, A.J. (2012) Genome-wide analyses of Zta binding to the Epstein-Barr virus genome reveals interactions in both early and late lytic cycles and an epigenetic switch leading to an altered binding profile. *J. Virol.*, **86**, 12494–12502.
- Flemington, E. and Speck, S.H. (1990) Epstein-Barr virus BZLF1 trans activator induces the promoter of a cellular cognate gene, c-fos. *J. Virol.*, **64**, 4549–4552.
- Heather, J., Flower, K., Isaac, S. and Sinclair, A.J. (2009) The Epstein-Barr virus lytic cycle activator Zta interacts with methylated ZRE in the promoter of host target gene egr1. *J. Gen. Virol.*, **90**, 1450–1454.
- Chang, Y., Lee, H.H., Chen, Y.T., Lu, J., Wu, S.Y., Chen, C.W., Takada, K. and Tsai, C.H. (2006) Induction of the early growth response 1 gene by Epstein-Barr virus lytic transactivator Zta. *J. Virol.*, **80**, 7748–7755.
- Hsu, M., Wu, S.Y., Chang, S.S., Su, I.J., Tsai, C.H., Lai, S.J., Shiao, A.L., Takada, K. and Chang, Y. (2008) Epstein-Barr virus lytic transactivator Zta enhances chemotactic activity through induction of interleukin-8 in nasopharyngeal carcinoma cells. *J. Virol.*, **82**, 3679–3688.
- Beatty, P.R., Krams, S.M. and Martinez, O.M. (1997) Involvement of IL-10 in the autonomous growth of EBV-transformed B cell lines. *J. Immunol.*, **158**, 4045–4051.
- Tsai, S.C., Lin, S.J., Chen, P.W., Luo, W.Y., Yeh, T.H., Wang, H.W., Chen, C.J. and Tsai, C.H. (2009) EBV Zta protein induces the expression of interleukin-13, promoting the proliferation of EBV-infected B cells and lymphoblastoid cell lines. *Blood*, **114**, 109–118.
- Yuan, J., Cahir-McFarland, E., Zhao, B. and Kieff, E. (2006) Virus and cell RNAs expressed during Epstein-Barr virus replication. *J. Virol.*, **80**, 2548–2565.
- Broderick, P., Hubank, M. and Sinclair, A.J. (2009) Effects of Epstein-Barr virus on host gene expression in Burkitt's lymphoma cell lines. *Chin. J. Cancer*, **28**, 813–821.
- Takada, K. (1984) Cross-linking of surface immunoglobulins induces Epstein-Barr virus in Burkitt's lymphoma cell lines. *Int. J. Cancer*, **33**, 27–32.
- Hollyoake, M., Stuhler, A., Farrell, P., Gordon, J. and Sinclair, A. (1995) The normal cell cycle activation program is exploited during the infection of quiescent B lymphocytes by Epstein-Barr virus. *Cancer Res.*, **55**, 4784–4787.
- Ben-Bassat, H., Goldblum, N., Mitrani, S., Klein, G. and Johansson, B. (1976) Concanavalin A receptors on the surface membrane of lymphocytes from patients with African Burkitt's lymphoma and lymphoma cell lines. *Int. J. Cancer*, **17**, 448–454.
- Zuo, J., Thomas, W.A., Haigh, T.A., Fitzsimmons, L., Long, H.M., Hislop, A.D., Taylor, G.S. and Rowe, M. (2011) Epstein-Barr virus evades CD4+ T cell responses in lytic cycle through BZLF1-mediated downregulation of CD74 and the cooperation of vBcl-2. *PLoS Pathog.*, **7**, e1002455.
- Young, L.S., Lau, R., Rowe, M., Niedobitek, G., Packham, G., Shanahan, F., Rowe, D.T., Greenspan, D., Greenspan, J.S., Rickinson, A.B. et al. (1991) Differentiation-associated expression of the Epstein-Barr virus BZLF1 transactivator protein in oral hairy leukoplakia. *J. Virol.*, **65**, 2868–2874.
- Klug, M. and Rehli, M. (2006) Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. *Epigenetics*, **1**, 127–130.
- Bailey, S.G., Verrall, E., Schelcher, C., Rhie, A., Doherty, A.J. and Sinclair, A.J. (2009) Functional interaction between Epstein-Barr virus replication protein Zta and host DNA-damage response protein 53BP1. *J. Virol.*, **83**, 11116–11122.
- Bark-Jones, S.J., Webb, H.M. and West, M.J. (2006) EBV EBNA 2 stimulates CDK9-dependent transcription and RNA polymerase II phosphorylation on serine 5. *Oncogene*, **25**, 1775–1785.
- Feng, J., Liu, T. and Zhang, Y. (2011) Using MACS to identify peaks from ChIP-Seq data. *Curr. Protoc. Bioinformatics*, Chapter 2, Unit 2.14.

38. Pruitt, K.D., Tatusova, T., Brown, G.R. and Maglott, D.R. (2012) NCBI reference sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.*, **40**, D130–D135.
39. Machanick, P. and Bailey, T.L. (2011) MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics*, **27**, 1696–1697.
40. Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L. and Noble, W.S. (2007) Quantifying similarity between motifs. *Genome Biol.*, **8**, R24.
41. Huang da, W., Sherman, B.T., Tan, Q., Collins, J.R., Alvord, W.G., Roayaei, J., Stephens, R., Baseler, M.W., Lane, H.C. and Lempicki, R.A. (2007) The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.*, **8**, R183.
42. Huang da, W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C. *et al.* (2007) DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.*, **35**, W169–W175.
43. Li, B. and Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323.
44. Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q. *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.*, **29**, 644–652.
45. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.*, **5**, R80.
46. Thorvaldsdottir, H., Robinson, J.T. and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinformatics*, **14**, 178–192.
47. Ramasubramanian, S., Osborn, K., Flower, K. and Sinclair, A.J. (2012) Dynamic chromatin environment of key lytic cycle regulatory regions of the Epstein-Barr virus genome. *J. Virol.*, **86**, 1809–1819.
48. Tsang, C.M., Zhang, G., Seto, E., Takada, K., Deng, W., Yip, Y.L., Man, C., Hau, P.M., Chen, H., Cao, Y. *et al.* (2010) Epstein-Barr virus infection in immortalized nasopharyngeal epithelial cells: regulation of infection and phenotypic characterization. *Int. J. Cancer*, **127**, 1570–1583.
49. Farnham, P.J. (2009) Insights from genomic profiling of transcription factors. *Nat. Rev. Genet.*, **10**, 605–616.
50. Krig, S.R., Jin, V.X., Bieda, M.C., O'Geen, H., Yaswen, P., Green, R. and Farnham, P.J. (2007) Identification of genes directly regulated by the oncogene ZNF217 using chromatin immunoprecipitation (ChIP)-chip assays. *J. Biol. Chem.*, **282**, 9703–9712.
51. Yang, A., Zhu, Z., Kapranov, P., McKeon, F., Church, G.M., Gingeras, T.R. and Struhl, K. (2006) Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. *Mol. Cell*, **24**, 593–602.
52. Scacheri, P.C., Davis, S., Odom, D.T., Crawford, G.E., Perkins, S., Halawi, M.J., Agarwal, S.K., Marx, S.J., Spiegel, A.M., Meltzer, P.S. *et al.* (2006) Genome-wide analysis of menin binding provides insights into MEN1 tumorigenesis. *PLoS Genet.*, **2**, e51.
53. Kenney, S.C. and Mertz, J.E. (2014) Regulation of the latent-lytic switch in Epstein-Barr virus. *Semin. Cancer Biol.*, **26**, 60–68.
54. Carey, M., Kolman, J., Katz, D.A., Gradoville, L., Barberis, L. and Miller, G. (1992) Transcriptional synergy by the Epstein-Barr virus transactivator ZEBRA. *J. Virol.*, **66**, 4803–4813.
55. Gustems, M., Woellmer, A., Rothbauer, U., Eck, S.H., Wieland, T., Lutter, D. and Hammerschmidt, W. (2014) c-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs. *Nucleic Acids Res.*, **42**, 3059–3072.
56. Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I. and Young, R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307–319.
57. Feederle, R., Mehl-Lautscham, A.M., Bannert, H. and Delecluse, H.J. (2009) The Epstein-Barr virus protein kinase BGLF4 and the exonuclease BGLF5 have opposite effects on the regulation of viral protein production. *J. Virol.*, **83**, 10877–10891.
58. Rowe, M., Glaunsinger, B., van Leeuwen, D., Zuo, J., Sweetman, D., Ganem, D., Middeldorp, J., Wiertz, E.J. and Rensing, M.E. (2007) Host shutoff during productive Epstein-Barr virus infection is mediated by BGLF5 and may contribute to immune evasion. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 3366–3371.
59. Zuo, J., Thomas, W., van Leeuwen, D., Middeldorp, J.M., Wiertz, E.J., Rensing, M.E. and Rowe, M. (2008) The DNase of gammaherpesviruses impairs recognition by virus-specific CD8+ T cells through an additional host shutoff function. *J. Virol.*, **82**, 2385–2393.
60. Inman, G.J., Binne, U.K., Parker, G.A., Farrell, P.J. and Allday, M.J. (2001) Activators of the Epstein-Barr virus lytic program concomitantly induce apoptosis, but lytic gene expression protects from cell death. *J. Virol.*, **75**, 2400–2410.
61. Chi, T., Lieberman, P., Lehman, A. and Carey, M. (1995) Mechanisms of transcriptional activation by zebra, an Epstein-Barr-virus protein. *FASEB J.*, **9**, A1462–A1462.
62. Lieberman, P.M. and Berk, A.J. (1994) A mechanism for TAFs in transcriptional activation: activation domain enhancement of TFIID-TFIIA-promoter DNA complex formation. *Genes Dev.*, **8**, 995–1006.
63. Lieberman, P. (1994) Identification of functional targets of the Zta transcriptional activator by formation of stable preinitiation complex intermediates. *Mol. Cell. Biol.*, **14**, 8365–8375.
64. Lieberman, P.M. and Berk, A.J. (1991) The Zta trans-activator protein stabilizes TFIID association with promoter DNA by direct protein-protein interaction. *Genes Dev.*, **5**, 2441–2454.
65. McClellan, M.J., Wood, C.D., Ojienyi, O., Cooper, T.J., Kanhere, A., Arvey, A., Webb, H.M., Palermo, R.D., Harth-Hertle, M.L., Kempkes, B. *et al.* (2013) Modulation of enhancer looping and differential gene targeting by Epstein-Barr virus transcription factors directs cellular reprogramming. *PLoS Pathog.*, **9**, e1003636.
66. Jiang, S., Willox, B., Zhou, H., Holthaus, A.M., Wang, A., Shi, T.T., Maruo, S., Kharchenko, P.V., Johannsen, E.C., Kieff, E. *et al.* (2014) Epstein-Barr virus nuclear antigen 3C binds to BATF/IRF4 or SPI1/IRF4 composite sites and recruits Sin3A to repress CDKN2A. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 421–426.

*Communication*

## Identification of Epstein-Barr Virus Replication Proteins in Burkitt's Lymphoma Cells

Chris Traylen <sup>1,†</sup>, Sharada Ramasubramanyan <sup>1,†</sup>, Jianmin Zuo <sup>2</sup>, Martin Rowe <sup>2</sup>,  
Rajaei Almohammad <sup>1</sup>, Kate Heesom <sup>4</sup>, Steve M. M. Sweet <sup>3</sup>, David A. Matthews <sup>5</sup> and  
Alison J. Sinclair <sup>1,\*</sup>

<sup>1</sup> School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK;

E-Mails: C.Traylen@sussex.ac.uk (C.T.); sramasubramanyan@gmail.com (S.R.);  
r.almohammed@sussex.ac.uk (R.A.)

<sup>2</sup> School of Cancer Sciences and Centre for Human Virology, University of Birmingham College of Medical and Dental Sciences, Edgbaston, Birmingham B15 2TT, UK;

E-Mails: j.zuo@bham.ac.uk (J.Z.); m.rowe@bham.ac.uk (M.R.)

<sup>3</sup> Genome Damage and Stability Centre, University of Sussex, Brighton BN1 9RQ, UK;

E-Mail: ss641@sussex.ac.uk

<sup>4</sup> University of Bristol Proteomics Facility, Biomedical Sciences Building, Bristol BS8 1TD, UK;

E-Mail: K.Heesom@bristol.ac.uk (K.H.)

<sup>5</sup> School of Cellular and Molecular Medicine, University of Bristol, Biomedical Sciences Building, Bristol BS8 1TD, UK; E-Mail: padam@bristol.ac.uk (D.A.M.)

<sup>†</sup> These authors contributed equally to this work.

\* Author to whom correspondence should be addressed; E-Mail: a.j.sinclair@sussex.ac.uk;  
Tel.: +44-1273-678-194; Fax: +44-1273-678-433.

Academic Editor: Lawrence S. Young

Received: 7 July 2015 / Accepted: 23 October 2015 / Published: 30 October 2015

---

**Abstract:** The working model to describe the mechanisms used to replicate the cancer-associated virus Epstein-Barr virus (EBV) is partly derived from comparisons with other members of the Herpes virus family. Many genes within the EBV genome are homologous across the herpes virus family. Published transcriptome data for the EBV genome during its lytic replication cycle show extensive transcription, but the identification of the proteins is limited. We have taken a global proteomics approach to identify viral proteins that are expressed during the EBV lytic replication cycle. We combined an



enrichment method to isolate cells undergoing EBV lytic replication with SILAC-labeling coupled to mass-spectrometry and identified viral and host proteins expressed during the EBV lytic replication cycle. Amongst the most frequently identified viral proteins are two components of the DNA replication machinery, the single strand DNA binding protein BALF2, DNA polymerase accessory protein BMRF1 and both subunits of the viral ribonucleoside-diphosphate reductase enzyme (BORF2 and BaRF1). An additional 42 EBV lytic cycle proteins were also detected. This provides proteomic identification for many EBV lytic replication cycle proteins and also identifies post-translational modifications.

**Keywords:** virus; cancer; replication; proteome; herpes; Epstein-Barr

---

## 1. Introduction

Epstein-Barr virus (EBV) is associated with diverse cancers including Burkitt's lymphoma, Hodgkin's lymphoma, NK/T lymphomas, Nasopharyngeal carcinoma and gastric cancer [1–9]. During the ~50-years since the identification of the virus [10] and the ~30 years since the genome sequence of the first isolate was published [11], there has been a strong focus on research into the viral genes commonly expressed in tumors, which has enabled us to obtain a good understanding of the ability of EBV to transform cells and so establish viral latency.

EBV within tumor cells undergoes lytic cycle replication only rarely and ~90% of EBV genes are not commonly expressed in tumors. However, these are transcribed following the disruption of latency as cells enter the EBV lytic replication cycle. Sensitive transcriptome analysis in Burkitt's lymphoma cells that have been stimulated to initiate the EBV lytic replication cycle [12,13], together with array-based strategies [14,15] and earlier mapping approaches (reviewed in [16]), suggests that the entire genome complement is expressed once EBV lytic replication cycle is activated.

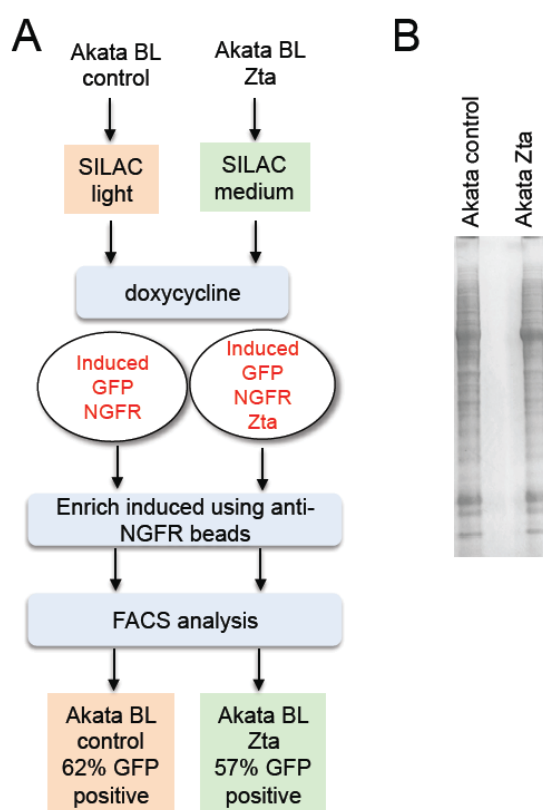
The contribution of several EBV lytic cycle genes has been subject to genetic evaluation. This identified BZLF1, BRLF1 [17], BSLF2 + BMLF1 [18] and BMRF1 [19] as essential for regulating viral gene expression during viral lytic replication and others (BFLF1, BFLF2, BFRF1, BGRF1 and BDRF1) contribute to encapsulating the viral genome [20–22]. In contrast, BGLF4 contributes to the efficiency of viral replication [23–26] and some viral genes such as BLLF1 and BNRF1 are not required to generate virus but rather contribute to the subsequent infection of cells or allow efficient entry and genome release [27–29]. Finally, some viral genes contribute to immune evasion of infected cells, e.g., BNLF2a [30]. The contributions that many other EBV lytic replication cycle genes make to the EBV lytic replication cycle are inferred through their homology with the alpha herpesvirus family [1]. Several of these proteins have been detected by immunofluorescence during viral replication (e.g., [31]).

Despite three studies using proteomics approaches [32–34], not all EBV lytic cycle genes have been previously identified and many have not been independently verified. Here, we used an engineered Burkitt's lymphoma cell system to enrich for cells undergoing EBV lytic replication and coupled this with SILAC-proteomics to develop a route to detect EBV proteins in Akata cells undergoing EBV lytic replication. This allowed us to identify a total of 44 EBV proteins and post-translational modifications of several viral proteins.

## 2. Results

### 2.1. Isolation of Proteins in Cells Undergoing EBV Lytic Cycle

Cells from a Burkitt's lymphoma which harbor EBV in type I latency had previously been engineered to co-express Green Fluorescent Protein (GFP), Nerve Growth Factor receptor (NGFR) and Zta (BZLF1) from an inducible bi-directional promoter (Akata-Zta). A cell line in which the Zta coding sequence is orientated in the non-coding direction acts as a control [35,36]. Proteins within the Akata control and Akata Zta cells were differentially metabolically labeled with amino acids consisting of stable isotopes. Following activation of the expression cassette using doxycycline, cells that had successfully been induced were isolated by their affinity for anti-NGFR coated magnetic beads. Analysis of GFP expression in the enriched cell population revealed a purity of between 57% and 62% (Figure 1).

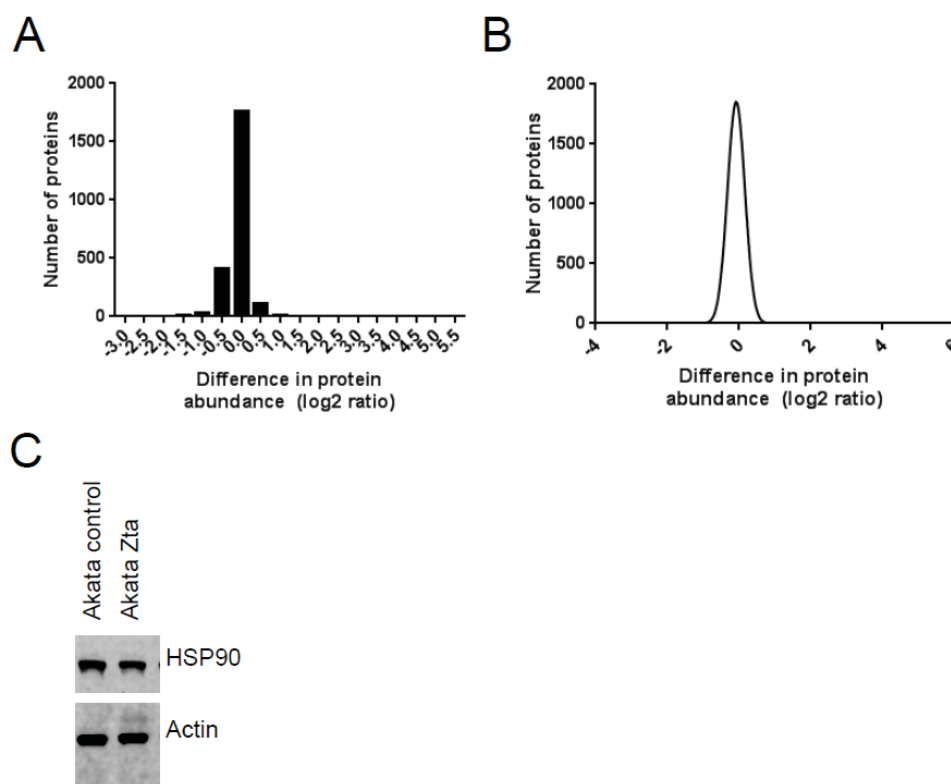


**Figure 1.** Enrichment of Burkitt's Lymphoma (BL) cells induced to enter Epstein-Barr virus (EBV) lytic replication cycle. (a) Co-induction of Green Fluorescent Protein (GFP), Nerve Growth factor receptor (NGFR) and Zta (or not for control cells) and procedure to induce and enrich cells, together with the % enrichment (GFP positivity) is shown; (b) Total protein extracts were prepared, fractionated on SDS-PAGE and stained.

### 2.2. Identification of Proteins in Cells Undergoing EBV Lytic Cycle

The proteins from Akata-control and Akata-Zta were mixed in equal amounts and the relative abundance of cellular and viral proteins was analyzed by quantitative mass spectrometry (MS). Global analysis of the differences in abundance detected through the differential SILAC labeling and the difference in abundance of individual proteins determined by Western blot analysis revealed a modest

overall reduction in the abundance of cellular proteins (between 1.2 and 2-fold) during EBV lytic cycle (Figure 2). Viral proteins were identified only in the Zta-expressing cells.



**Figure 2.** Stable isotope labeling with amino acids in cell culture (SILAC) coupled to mass spectrometry (MS) analysis of proteins in Burkitt's Lymphoma (BL) cells during EBV lytic cycle. Total protein extracts were prepared from the enriched BL cells. MS analysis was undertaken. (a) The change in abundance of proteins with SILAC-information from both control and Zta-expressing cells is shown; (b) The frequency distribution of the difference in protein abundance is shown as a Gaussian plot; (c) Total proteins were separated by SDS-PAGE. Western blots were probed with anti-HSP90 and beta actin antibodies.

### 2.3. Identification of EBV Proteins

To identify EBV proteins in cells undergoing lytic replication, we considered the peptides that match with an EBV protein. The Uniprot databases which include proteins from three viral genomes HHV4 (B95-8 UP000007640; AG876 UP000007639; and GD1 UP000007641). The identity of each of the 169 peptides that correspond to an EBV protein with a Posterior Error Probability (PEP) score of less than  $1.0 \times 10^{-3}$  are provided in Table S1. These were all up regulated  $\geq 8.6$  fold during lytic cycle, with the majority being undetectable in latency. This identified peptides corresponding to 33 EBV proteins (Table 1). In addition, a custom made database of the Akata EBV proteome was generated and searched to ensure that polymorphic regions were not overlooked. However, this revealed no additional protein identifications. In order to increase the sensitivity of EBV protein detection in our dataset, we carried out a further MaxQuant search against the EBV proteome (UniProt), omitting human sequences [37]. This identified an additional 11 EBV proteins, also shown in Table 1 (highlighted). The peptides associated with this search are listed in Table S2.

**Table 1.** EBV proteins identified by MS analysis.

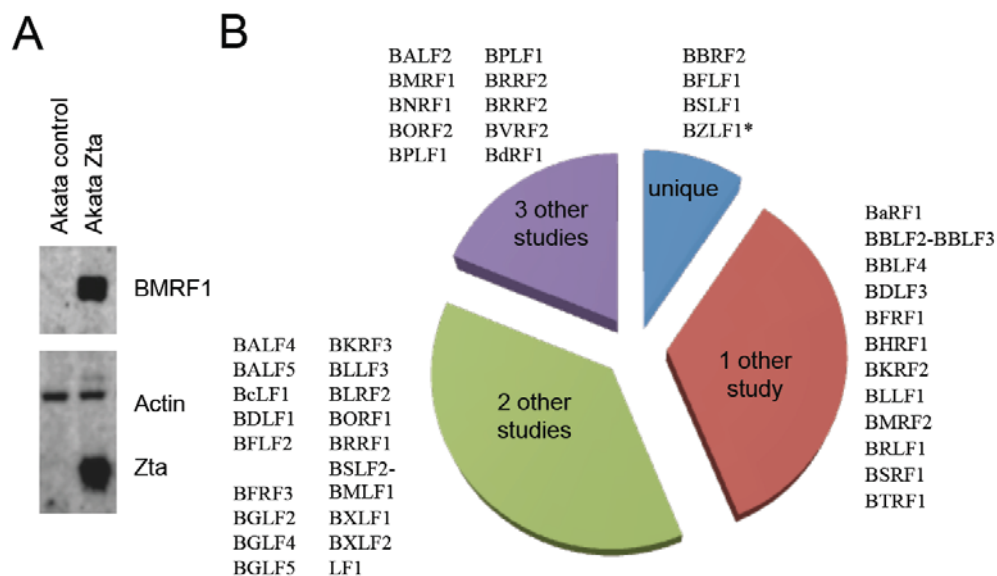
Gene	Function
BALF2	Major DNA-binding protein
BALF4	Envelope glycoprotein B
BALF5	DNA polymerase catalytic subunit
BaRF1	Ribonucleoside-diphosphate reductase small chain
BBLF2-BBLF3	primase protein
BBLF4	DNA replication helicase
BBRF2	Virion egress protein UL7 homolog
BcLF1	Major capsid protein
BDLF1	Triplex capsid protein VP23 homolog
BFLF1	Packaging protein UL32 homolog
BFLF2	Virion egress protein
BFRF1	Virion egress protein UL34 homolog
BFRF3	Capsid protein VP26
BGLF2	Capsid-binding protein
BGLF4	Serine/threonine-protein kinase
BGLF5	Shutoff alkaline exonuclease
BHRF1	Apoptosis regulator
BKRF3	Uracil-DNA glycosylase
BLLF3	Deoxyuridine 5'-triphosphate nucleotidohydrolase
BLRF2	Tegument protein
BSLF2-BMLF1	mRNA export factor ICP27 homolog
BMRF1	DNA polymerase processivity factor
BNRF1	Major tegument protein
BORF2	Ribonucleoside-diphosphate reductase large subunit
BPLF1	Deneddylase
BRRF1	Transcriptional activator
BRRF2	Tegument protein
BSRF1	Tegument protein UL51 homolog
BTRF1	Uncharacterized protein BTRF1
BVRF2	Capsid scaffolding protein
BdRF1	
BXLF1	Thymidine kinase
BZLF1 *	Trans-activator protein
BDLF3	pg85
BLLF1	gp350
BMRF2	Protein BMRF2
BORF1	Triplex capsid protein
BPLF1	deneddylase
BRLF1	Replication and transcription factor
BRRF2	tegument protein
BSLF1	DNA primase
gH	gH
gL	gL
LF1	LF1

\* BZLF1 expression is driven by the doxycycline induced expression vector in these cells so detection cannot be ascribed to the endogenous protein. Yellow highlight represents proteins only identified in the EBV-specific search.

**Table 2.** Post-translational modifications of EBV proteins identified by MS analysis.

Gene Name	Modification	pep_seq	aa of EBV Protein	Residue of Modification
BALF5	N terminal acetylation	[ac]SGGLFYNPFLRPNK	2–15	2
BLLF3	N terminal acetylation	[ac]MEACPHIR	9–16	9
BLRF2	Phosphorylation	GQPS[ph]PGEGRPR	124–135	127
BMRF1	2 Phosphorylation	HTVS[ph]PSPS[ph]PPPPPR	330–343	333 and 337
BMRF2	N terminal acetylation	[ac]METTQTLR	1–8	1
BORF1	Phosphorylation	RLNIS[ph]R	26–31	30
BORF2	N terminal acetylation	[ac]ATTSHVEHELLSK	2–14	2
BXLF1	Phosphorylation	TQAAVTSNTGNS[ph]PGSR	86–101	97
BZLF1	N terminal acetylation	[ac]MMDPNSTSEDVK	1–12	1

None of the EBV proteins are associated with EBV latency; all originate from genes with a characteristic lytic replication cycle pattern of expression [16]. One of the 44 proteins identified, BZLF1, could be derived from either the expression vector or the endogenous virus so it should not be considered as proof of identity of the endogenous protein. We confirm expression of one of these proteins, BMRF1, by Western blot (Figure 3A) and we show which gene products are uniquely identified here and which provides confirmation of proteins previously identified in other reports (Figure 3B).



**Figure 3.** SILAC MS analysis of proteins EBV proteins detected in Akata cells during lytic cycle. (A) Akata control and Akata Zta cells were induced with doxycycline for 24 h and total protein extracts prepared. Western blot analysis of BMRF1, Zta and beta actin abundance is show; (B) The EBV proteins identified are shown in relation to previously published studies. BZLF1 is marked\*, its expression is driven by the doxycycline induced expression vector in these cells so detection cannot be ascribed to the endogenous protein.

#### 2.4. Identification of Post-Translational Modifications of EBV Proteins

We searched for potential post-translational modifications of EBV proteins using MASCOT to search a minimal EBV database. All identifications were required to be from medium-labeled peptides, *i.e.*,

present after EBV induction. The same phosphorylation and N-terminal acetylation modifications were also identified in a MaxQuant search of the EBV database. Those phospho-serine and amino-terminal acetylation modifications corresponding to proteins identified in Table 1 are shown in Table 2, with peptide identification evidence provided in Table S3. This identified that five lytic EBV proteins sustain amino-terminal acetylation; BZLF1 (Zta), BMRF2, BLLF3, BALF5, and BORF2. In addition EBV peptides corresponding to serine phosphorylation were identified for BMRF1, BLRF2, BORF1 and BXLF1.

### 3. Experimental Section

#### 3.1. Cell Culture

Akata-Zta and Akata control cells [36] were cultured in RPMI–SILAC labeled RPMI containing 13C6-arginine and 4,4,5,5-D4-lysine (R6K4) (medium) and RPMI R0K0 (light) respectively (Dundee Cell products). Each was supplemented with 15% (v/v) dialyzed FBS and 100 units/mL penicillin, 100 µg/mL streptomycin and 2 mM L-glutamine (Life Technologies) at 37 °C with 5% CO<sub>2</sub>. Cells were maintained between 3 and 10 × 10<sup>5</sup> cells/mL and were cultured in SILAC-medium for 16 population doublings. Doxycycline (Sigma) was added to a final concentration of 500 ng/mL and cells incubated for a further 24 h. Successfully induced cells were enriched using anti-NGFR antibodies coupled to paramagnetic beads as described [36].

#### 3.2. FACS Analysis

Live cells were analyzed using a multi-parameter fluorescent activated cell analysis (FACs) (Facs Canto-Beckton Dickinson). GFP positive cells were identified using BD FACSDiva™ Software (Beckton Dickinson).

#### 3.3. Western Blot Analysis

An equivalent number of cells were lysed using SDS-PAGE sample buffer at a final concentration of 1.0 × 10<sup>4</sup> cells/µL. Extracts from 1.0 × 10<sup>5</sup> cells were fractionated on SDS-PAGE. The total protein complement was detected following staining with Simply Blue stain or transferred to nitrocellulose. Proteins were identified using the following primary antibodies, beta actin (A2066, SIGMA), HSP90 (AC88 ab13492, ABCAM), Zta [38] and BMRF1 (8F92, ab30541, ABCAM). This was followed by incubation with species-specific infra-red labeled secondary antibodies (LiCor). The presence and relative abundance of proteins was determined using an Odyssey Fc Imager and Odyssey Image Studio (Licor).

#### 3.4. Mass Spectrometry Collection and Analysis

Extracts from Zta expressing and not expressing cells were mixed and fractionated on SDS-PAGE (Novex). The lane was cut into six slices and each slice was subjected to in-gel digestion with a DigestPro MSi automatic digestion system (Intavis Bioanalytical Instruments) as described in [39]. The resulting peptides were fractionated using a Dionex Ultimate 3000 nano HPLC system coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). In brief, peptides in 1% (v/v) formic acid were injected onto an Acclaim PepMap C18 nano-trap column (Dionex). After washing with 0.5% (v/v)



acetonitrile 0.1% (v/v) formic acid, peptides were resolved on a 250 mm × 75 µm Acclaim PepMap C18 reverse phase analytical column (Dionex) over a 150 min organic gradient, using 7 gradient segments (1%–6% solvent B over 1 min, 6%–15% B over 58 min, 15%–32% B over 58 min, 32%–40% B over 3 min, 40%–90% B over 1 min, held at 90% B for 6 min and then reduced to 1% B over 1 min) with a flow rate of 300 nL·min<sup>−1</sup>. Solvent A was 0.1% formic acid and Solvent B was aqueous 80% acetonitrile in 0.1% formic acid. Peptides were ionized by nano-electrospray ionization at 2.3 kV using a stainless steel emitter with an internal diameter of 30 µm (Thermo Scientific) and a capillary temperature of 250 °C. Tandem mass spectra were acquired using an LTQ-Orbitrap Velos mass spectrometer controlled by Xcalibur v2.1 software [40] and operated in data-dependent acquisition mode. The Orbitrap was set to analyze the survey scans at 60,000 resolution (at m/z 400) in the mass range m/z 300–2000 and the top six multiply charged ions in each duty cycle selected for MS/MS in the LTQ linear ion trap. Charge state filtering, where unassigned precursor ions were not selected for fragmentation, and dynamic exclusion (repeat count, 1; repeat duration, 30 s; exclusion list size, 500) were used. Fragmentation conditions in the LTQ were as follows: normalized collision energy, 40%; activation q, 0.25; activation time 10 ms; and minimum ion selection intensity, 500 counts. Data were acquired using the Xcalibur v2.1 software. The raw data files were processed and quantified using MaxQuant as described in [39] and searched against standard human proteome and EBV protein lists from UNIPROT and a translation of the Akata EBV genome. A search was also carried out against the EBV UniProt proteins plus contaminants, with the human sequences omitted [37]. Peptide precursor mass tolerance was set at 10 ppm, and MS/MS tolerance was set at 0.8 Da. Search criteria included carbamidomethylation of cysteine (+57.0214) as a fixed modification and oxidation of methionine (+15.9949) and appropriate SILAC labels as variable modifications.

Searches were performed with full tryptic digestion and a maximum of two missed cleavages was allowed. The reverse database search option was enabled and all peptide data was filtered to satisfy false discovery rate (FDR) of 1%.

A database search using Mascot was carried out against a database containing 282 EBV protein sequences from UniProt. The search used the following parameters: 10 ppm precursor mass tolerance; 0.6 Da fragment ion mass tolerance; fixed modification: carbamidomethylation (C); variable modifications: Protein N-terminus acetylation, methionine oxidation, phosphorylation (STY), 2H(4) K, 13C(6) R.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium [1] via the PRIDE partner repository with the dataset identifier PXD002461 [41].

#### 4. Conclusions

Two previous studies compared the proteomes in BL and primary effusion lymphoma (PEL) cells during EBV lytic cycle with the proteomes of cells during latency or to those that are refractory to entering EBV lytic cycle [33,34]. The previous studies used the histone deacetylase inhibitor sodium butyrate [34] and/or a combination of the histone deacetylase inhibitor sodium butyrate and 12-O-tetradecanoylphorbol-13-acetate to induce EBV to enter its lytic replication cycle. Here, we used a different method to initiate EBV lytic cycle gene expression, the ectopic expression of Zta protein. We previously demonstrated that this is sufficient to promote expression of several EBV lytic cycle genes leading to replication of the EBV genome [36]. Sensitive transcriptome analysis of EBV identified highly

abundant mRNAs [12,14,33]. While some of the proteins encoded by these are readily detected in the lytic cells (e.g., BMRF1, BMLF1 and BHRF1), others are not detected in any study (e.g., BALF1). This highlights one of the limitations of interpreting a global proteomics study; some proteins do not generate peptides that can be unambiguously identified. Whether EBV completes the lytic cycle in response to any of these stimuli yet protein expression is too low to be detected by mass spectrometry, or whether the lytic cycles are aborted prior to full viral gene expression and release of infectious virus is unknown. In addition to the analysis of viral proteins within cells, proteins present in purified EBV virions have also been detected using proteomics [32].

A comparison of our data with these three datasets revealed that we detected 28 EBV proteins that had been identified in two or more previous studies. We therefore provide further support for the identification of these proteins. Importantly, our analysis detected 12 viral proteins that were only identified in one previous study, providing important independent evidence of their detection. In addition, we provide evidence for the first detection of three viral proteins by mass spectrometry. The first is BBRF2, which is a homologue of the HSV1 virion egress protein UL7. Clues as to its function arise from the recent demonstration that UL7 plays a role in linking tegument proteins of HSV1 to membranes [42]. The second protein is BFLF1. Interestingly, BFLF1 is a homologue of the HSV1 UL32 gene, which plays a role in HSV1 encapsidation [43] which supports the potential involvement of BFLF1 protein in cleavage and packaging of the viral genome [21]. The third, BSLF1, encodes the DNA primase that is required for genome lytic replication [44,45]. In addition, we detected Zta protein (BZLF1), although it is not possible to distinguish whether this originates from the endogenous genome or the expression vector.

Further analysis of the data identified evidence for novel post-translational modifications of nine EBV proteins. Amino-terminal acetylation events were identified for Zta, BLLF3, BALF5, BMRF2 and BORF2. For BALF5 and BORF2 amino terminal processing had occurred and the acetylation is present on the second residue, for the remainder it is present on the initiator methionine. Neither the acetylation nor the amino terminal processing had been described previously. A large sub-set of cellular proteins also sustain the amino terminal acetylation, the function is enigmatic, and roles in protein–protein interaction, sub-cellular targeting and degradation have all been postulated [46]. In addition to this, EBV peptides corresponding to serine phosphorylation of BMRF1, BLRF2, BORF1 and BXLF1 were identified. Of these, BMRF1 is known to be phosphorylated at residue 337 [47], in addition to residues 344, 349 and 355. We provide evidence for a further site of phosphorylation at serine 333. In addition, this is the first report that BLRF2, BORF2 and BXLF1 sustain serine phosphorylation.

In summary, the definitive identification of 44 EBV proteins in BL cells undergoing EBV replication and the identification of novel post-translational modifications of nine of these lytic cycle proteins increase the knowledge base of EBV lytic replication and may highlight different targets for future strategies to enable the development of therapeutic interventions to manipulate EBV replication.

## Supplementary Materials

Supplementary materials can be accessed at: <http://www.mdpi.com/2076-0817/4/3/739/s1>.



## Acknowledgments

The research was funded by grants from the Medical Research Council (MR/J001708/1 to AJS and MR, G0901755 to MR) and from the BBSRC (BB/L018438/1 to DAM).

## Author Contributions

Martin Rowe and Jianmin Zuo established the cell lines used in the study. Chris Traylen, Sharada Ramasubramanian, Rajaei Almohammad undertook SILAC labeling, cell purification and mass spec analysis. Kate Heesom, David A. Matthews, Steve M. M. Sweet, Chris Traylen and Alison J. Sinclair analyzed data. Alison J. Sinclair prepared the report.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Longnecker, R.; Kieff, E.; Cohen, J.I. Epstein-barr virus/replication and Epstein-Barr virus. In *Fields Virology*, 6th ed.; Lippencott, Williams, Wilkins: Baltimore, MD, USA, 2013.
2. Rowe, M.; Kelly, G.L.; Bell, A.I.; Rickinson, A.B. Burkitt's lymphoma: The rosetta stone deciphering Epstein-Barr virus biology. *Semin. Cancer Biol.* **2009**, *19*, 377–388.
3. Strong, M.J.; Xu, G.; Coco, J.; Baribault, C.; Vinay, D.S.; Lacey, M.R.; Strong, A.L.; Lehman, T.A.; Seddon, M.B.; Lin, Z.; *et al.* Differences in gastric carcinoma microenvironment stratify according to EBV infection intensity: Implications for possible immune adjuvant therapy. *PLoS Pathog.* **2013**, *9*, e1003341.
4. Chen, X.Z.; Chen, H.; Castro, F.A.; Hu, J.K.; Brenner, H. Epstein-Barr virus infection and gastric cancer: A systematic review. *Medicine* **2015**, *94*, e792.
5. Shinozaki-Ushiku, A.; Kunita, A.; Fukayama, M. Update on Epstein-Barr virus and gastric cancer (review). *Int. J. Oncol.* **2015**, *46*, 1421–1434.
6. Kuppers, R. The biology of Hodgkin's lymphoma. *Nat. Rev. Cancer* **2009**, *9*, 15–27.
7. Vockerodt, M.; Cader, F.Z.; Shannon-Lowe, C.; Murray, P. Epstein-barr virus and the origin of Hodgkin lymphoma. *Chin. J. Cancer* **2014**, *33*, 591–597.
8. Lung, M.A. Acute effects of inhaled sulphur dioxide on pig nasal vascular and airway resistances. *Acta physiol. Sinica* **2014**, *66*, 79–84.
9. Chan, A.T. Nasopharyngeal carcinoma. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* **2010**, *21*, vii308–vii312.
10. Epstein, M.A.; Achong, B.G.; Barr, Y.M. Virus particles in cultured lymphoblasts from Burkitt's lymphoma. *Lancet* **1964**, *1*, 702–703.
11. Baer, R.; Bankier, A.T.; Biggin, M.D.; Deninger, P.D.; Farrell, P.J.; Gibson, T.J.; Hatfull, G.; Hudson, G.S.; Satchwell, S.C.; Seguin, C.; *et al.* DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature* **1984**, *310*, 207–211.

12. O'Grady, T.; Cao, S.; Strong, M.J.; Concha, M.; Wang, X.; Splinter Bondurant, S.; Adams, M.; Baddoo, M.; Srivastav, S.K.; Lin, Z.; *et al.* Global bidirectional transcription of the Epstein-Barr virus genome during reactivation. *J. Virol.* **2014**, *88*, 1604–1616.
13. Concha, M.; Wang, X.; Cao, S.; Baddoo, M.; Fewell, C.; Lin, Z.; Hulme, W.; Hedges, D.; McBride, J.; Flemington, E.K. Identification of new viral genes and transcript isoforms during Epstein-Barr virus reactivation using RNA-Seq. *J. Virol.* **2012**, *86*, 1458–1467.
14. Tierney, R.J.; Shannon-Lowe, C.D.; Fitzsimmons, L.; Bell, A.I.; Rowe, M. Unexpected patterns of Epstein-Barr virus transcription revealed by a high throughput PCR array for absolute quantification of viral mRNA. *Virology* **2015**, *474*, 117–130.
15. Kurokawa, M.; Ghosh, S.K.; Ramos, J.C.; Mian, A.M.; Toomey, N.L.; Cabral, L.; Whitby, D.; Barber, G.N.; Dittmer, D.P.; Harrington, W.J., Jr. Azidothymidine inhibits nf-kappaB and induces Epstein-Barr virus gene expression in Burkitt lymphoma. *Blood* **2005**, *106*, 235–240.
16. Farrell, P.J. Epstein-barr virus genome. In *Epstein-Barr Virus*; Robertson, E.S., Ed.; Caister: Wymondham, UK, 2005; pp. 263–288.
17. Feederle, R.; Kost, M.; Baumann, M.; Janz, A.; Drouet, E.; Hammerschmidt, W.; Delecluse, H.J. The Epstein-Barr virus lytic program is controlled by the co-operative functions of two transactivators. *EMBO J.* **2000**, *19*, 3080–3089.
18. Gruffat, H.; Batisse, J.; Pich, D.; Neuhierl, B.; Manet, E.; Hammerschmidt, W.; Sergeant, A. Epstein-Barr virus mRNA export factor eb2 is essential for production of infectious virus. *J. Virol.* **2002**, *76*, 9635–9644.
19. Neuhierl, B.; Delecluse, H.J. The Epstein-Barr virus bmf1 gene is essential for lytic virus replication. *J. Virol.* **2006**, *80*, 5078–5081.
20. Farina, A.; Feederle, R.; Raffa, S.; Gonnella, R.; Santarelli, R.; Frati, L.; Angeloni, A.; Torrisi, M.R.; Faggioni, A.; Delecluse, H.J. Bfrf1 of Epstein-Barr virus is essential for efficient primary viral envelopment and egress. *J. Virol.* **2005**, *79*, 3703–3712.
21. Granato, M.; Feederle, R.; Farina, A.; Gonnella, R.; Santarelli, R.; Hub, B.; Faggioni, A.; Delecluse, H.J. Deletion of Epstein-Barr virus bflf2 leads to impaired viral DNA packaging and primary egress as well as to the production of defective viral particles. *J. Virol.* **2008**, *82*, 4042–4051.
22. Pavlova, S.; Feederle, R.; Gartner, K.; Fuchs, W.; Granzow, H.; Delecluse, H.J. An Epstein-Barr virus mutant produces immunogenic defective particles devoid of viral DNA. *J. Virol.* **2013**, *87*, 2011–2022.
23. Murata, T.; Isomura, H.; Yamashita, Y.; Toyama, S.; Sato, Y.; Nakayama, S.; Kudoh, A.; Iwahori, S.; Kanda, T.; Tsurumi, T. Efficient production of infectious viruses requires enzymatic activity of Epstein-Barr virus protein kinase. *Virology* **2009**, *389*, 75–81.
24. El-Guindy, A.; Lopez-Giraldez, F.; Delecluse, H.J.; McKenzie, J.; Miller, G. A locus encompassing the Epstein-Barr virus bglf4 kinase regulates expression of genes encoding viral structural proteins. *PLoS Pathog.* **2014**, *10*, e1004307.
25. Feederle, R.; Bannert, H.; Lips, H.; Muller-Lantzsch, N.; Delecluse, H.J. The Epstein-Barr virus alkaline exonuclease bglf5 serves pleiotropic functions in virus replication. *J. Virol.* **2009**, *83*, 4952–4962.

26. Feederle, R.; Mehl-Lautscham, A.M.; Bannert, H.; Delecluse, H.J. The Epstein-Barr virus protein kinase bglf4 and the exonuclease bglf5 have opposite effects on the regulation of viral protein production. *J. Virol.* **2009**, *83*, 10877–10891.
27. Janz, A.; Oezel, M.; Kurzeder, C.; Mautner, J.; Pich, D.; Kost, M.; Hammerschmidt, W.; Delecluse, H.J. Infectious Epstein-Barr virus lacking major glycoprotein bllf1 (gp350/220) demonstrates the existence of additional viral ligands. *J. Virol.* **2000**, *74*, 10142–10152.
28. Feederle, R.; Neuhierl, B.; Baldwin, G.; Bannert, H.; Hub, B.; Mautner, J.; Behrends, U.; Delecluse, H.J. Epstein-Barr virus bnrf1 protein allows efficient transfer from the endosomal compartment to the nucleus of primary B lymphocytes. *J. Virol.* **2006**, *80*, 9435–9443.
29. Neuhierl, B.; Feederle, R.; Adhikary, D.; Hub, B.; Geletneky, K.; Mautner, J.; Delecluse, H.J. Primary B-cell infection with a deltabalf4 Epstein-Barr virus comes to a halt in the endosomal compartment yet still elicits a potent cd4-positive cytotoxic T-cell response. *J. Virol.* **2009**, *83*, 4616–4623.
30. Croft, N.P.; Shannon-Lowe, C.; Bell, A.I.; Horst, D.; Kremmer, E.; Rensing, M.E.; Wiertz, E.J.; Middeldorp, J.M.; Rowe, M.; Rickinson, A.B.; *et al.* Stage-specific inhibition of mhc class i presentation by the Epstein-Barr virus bnlf2a protein during virus lytic cycle. *PLoS Pathog.* **2009**, *5*, e1000490.
31. Sato, Y.; Tsurumi, T. Genome guardian p53 and viral infections. *Rev. Med. Virol.* **2013**, *23*, 213–220.
32. Johannsen, E.; Luftig, M.; Chase, M.R.; Weicksel, S.; Cahir-McFarland, E.; Illanes, D.; Sarracino, D.; Kieff, E. Proteins of purified Epstein-Barr virus. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16286–16291.
33. Dresang, L.R.; Teuton, J.R.; Feng, H.; Jacobs, J.M.; Camp, D.G., 2nd; Purvine, S.O.; Gritsenko, M.A.; Li, Z.; Smith, R.D.; Sugden, B.; *et al.* Coupled transcriptome and proteome analysis of human lymphotropic tumor viruses: Insights on the detection and discovery of viral genes. *BMC Genomics* **2011**, *12*, 625.
34. Koganti, S.; Clark, C.; Zhi, J.; Li, X.; Chen, E.I.; Chakraborty, S.; Hill, E.R.; Bhaduri-McIntosh, S. Cellular stat3 functions via pcbp2 to restrain ebv lytic activation in B lymphocytes. *J. Virol.* **2015**, *89*, 5002–5011.
35. Zuo, J.M.; Thomas, W.A.; Haigh, T.A.; Fitzsimmons, L.; Long, H.M.; Hislop, A.D.; Taylor, G.S.; Rowe, M. Epstein-barr virus evades cd4(+) t cell responses in lytic cycle through bzlf1-mediated downregulation of cd74 and the cooperation of vbcl-2. *PLoS Pathog.* **2011**, *7*, e1002455.
36. Ramasubramanian, S.; Osborn, K.; Al-Mohammad, R.; Naranjo Perez-Fernandez, I.B.; Zuo, J.; Balan, N.; Godfrey, A.; Patel, H.; Peters, G.; Rowe, M.; *et al.* Epstein-Barr virus transcription factor zta acts through distal regulatory elements to directly control cellular gene expression. *Nucleic Acids Res.* **2015**, *43*, 3563–3577.
37. Noble, W.S. Mass spectrometrists should search only for peptides they care about. *Nat. Methods* **2015**, *12*, 605–608.
38. Young, L.S.; Lau, R.; Rowe, M.; Niedobitek, G.; Packham, G.; Shanahan, F.; Rowe, D.T.; Greenspan, D.; Greenspan, J.S.; Rickinson, A.B.; *et al.* Differentiation-associated expression of the Epstein-Barr virus bzlf1 transactivator protein in oral hairy leukoplakia. *J. Virol.* **1991**, *65*, 2868–2874.

39. Evans, V.C.; Barker, G.; Heesom, K.J.; Fan, J.; Bessant, C.; Matthews, D.A. De novo derivation of proteomes from transcriptomes for transcript and protein identification. *Nat. Methods* **2012**, *9*, 1207–1211.
40. Thermo Scientific. Xcalibur Software v2.1. Available online: <http://www.thermoscientific.com/content/tfs/en/product/xcalibur-software.html> (accessed on 23 October 2015).
41. Vizcaino, J.A.; Deutsch, E.W.; Wang, R.; Csordas, A.; Reisinger, F.; Rios, D.; Dianes, J.A.; Sun, Z.; Farrah, T.; Bandeira, N.; *et al.* Proteomexchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **2014**, *32*, 223–226.
42. Roller, R.J.; Fetters, R. The herpes simplex virus 1 ul51 protein interacts with the ul7 protein and plays a role in its recruitment into the virion. *J. Virol.* **2015**, *89*, 3112–3122.
43. Lamberti, C.; Weller, S.K. The herpes simplex virus type 1 cleavage/packaging protein, ul32, is involved in efficient localization of capsids to replication compartments. *J. Virol.* **1998**, *72*, 2463–2473.
44. Fixman, E.D.; Hayward, G.S.; Hayward, S.D. Trans-acting requirements for replication of Epstein-Barr virus ori-lyt. *J. Virol.* **1992**, *66*, 5030–5039.
45. Fixman, E.D.; Hayward, G.S.; Hayward, S.D. Replication of Epstein-Barr virus orilyt: Lack of a dedicated virally encoded origin-binding protein and dependence on Zta in cotransfection assays. *J. Virol.* **1995**, *69*, 2998–3006.
46. Arnesen, T. Towards a functional understanding of protein N-terminal acetylation. *PLoS Biol.* **2011**, *9*, e1001074.
47. Yang, P.W.; Chang, S.S.; Tsai, C.H.; Chao, Y.H.; Chen, M.R. Effect of phosphorylation on the transactivation activity of Epstein-Barr virus bmrfl, a major target of the viral bglf4 kinase. *J. Gen. Virol.* **2008**, *89*, 884–895.